0-315-03566-8

National Library of Canada | Bibliothèque nationale du Canada

Canadian Theses Division | Division des thèses canadiennes

Ottawa, Canada
K1A 0N4

51461

PERMISSION TO MICROFILM — AUTORISATION DE MICROFILMER

• Please print or type — Écrire en lettres moulées ou dactylographier

Full Name of Author — Nom complet de l'auteur

CATHERINE VERONICA DESCHENEAU

Date of Birth — Date de naissance

6 Feb 1948

Country of Birth — Lieu de naissance

SOUTH AFRICA

Permanent Address — Résidence fixe

11136 49 AVE EDMONTON ALTA T6H 0G6

Title of Thesis — Titre de la thèse

MODELLING AND SIMULATION IN DEVELOPMENTAL GENETICS

University — Université

UNIVERSITY OF ALBERTA

Degree for which thesis was presented — Grade pour lequel cette thèse fut présentée

DOCTOR OF PHILOSOPHY

Year this degree conferred — Année d'obtention de ce grade

1981

Name of Supervisor — Nom du directeur de thèse

JEFFREY R. SAMPSON

Permission is hereby granted to the NATIONAL LIBRARY OF CANADA to microfilm this thesis and to lend or sell copies of the film.

The author reserves other publication rights, and neither the thesis nor extensive extracts from it may be printed or otherwise reproduced without the author's written permission.

L'autorisation est, par la présente, accordée à la BIBLIOTHÈQUE NATIONALE DU CANADA de microfilmer cette thèse et de prêter ou de vendre des exemplaires du film.

L'auteur se réserve les autres droits de publication; ni la thèse ni de longs extraits de celle-ci ne doivent être imprimés ou autrement reproduits sans l'autorisation écrite de l'auteur.

Date

15th April 1981

Signature

C. Descheneau

NL-91 (4/77)

National Library of Canada
Collections Development Branch

Canadian Theses on
Microfiche Service

Bibliothèque nationale du Canada
Direction du développement des collections

Service des thèses canadiennes
sur microfiche

NOTICE

The quality of this microfiche is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us a poor photocopy.

Previously copyrighted materials (journal articles, published tests, etc.) are not filmed.

Reproduction in full or in part of this film is governed by the Canadian Copyright Act, R.S.C. 1970, c. C-30. Please read the authorization forms which accompany this thesis.

AVIS

La qualité de cette microfiche dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade.

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de mauvaise qualité.

Les documents qui font déjà l'objet d'un droit d'auteur (articles de revue, examens publiés, etc.) ne sont pas microfilmés.

La reproduction, même partielle, de ce microfilm est soumise à la Loi canadienne sur le droit d'auteur, SRC 1970, c. C-30. Veuillez prendre connaissance des formules d'autorisation qui accompagnent cette thèse.

THIS DISSERTATION
HAS BEEN MICROFILMED
EXACTLY AS RECEIVED

LA THÈSE A ÉTÉ
MICROFILMÉE TELLE QUE
NOUS L'AVONS REÇUE

Ottawa, Canada
K1A 0N4

NL-339 (Rev. 8/80)

THE UNIVERSITY OF ALBERTA

MODELLING AND SIMULATION IN DEVELOPMENTAL GENETICS

by

© CATHERINE DESCHENEAU

University of Alberta

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES AND RESEARCH

IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE

OF DOCTOR OF PHILOSOPHY

DEPARTMENT OF COMPUTING SCIENCE

EDMONTON, ALBERTA

SPRING, 1981

THE UNIVERSITY OF ALBERTA

RELEASE FORM

NAME OF AUTHOR Catherine Descheneau

TITLE OF THESIS Modelling and Simulation in Developmental Genetics

DEGREE FOR WHICH THESIS WAS PRESENTED Doctor of Philosophy

YEAR THIS DEGREE GRANTED 1981

Permission is hereby granted to THE UNIVERSITY OF ALBERTA LIBRARY to reproduce single copies of this thesis and to lend or sell such copies for private, scholarly or scientific research purposes only.

The author reserves other publication rights, and neither the thesis nor extensive extracts from if may be printed or otherwise reproduced without the author's written permission.

(Signed) Catherine Descheneau

PERMANENT ADDRESS:

11136 - 49 Avenue

Edmonton, Alberta, Canada

DATED 15 April 1981.

University of Alberta

THE UNIVERSITY OF ALBERTA

FACULTY OF GRADUATE STUDIES AND RESEARCH

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies and Research, for acceptance, a thesis entitled MODELLING AND SIMULATION IN DEVELOPMENTAL GENETICS submitted by CATHERINE DESCHENEAU in partial fulfilment of the requirements for the degree of DOCTOR OF PHILOSOPHY.

.....

Supervisor

.....

Kenneth Morgan

.....

.....

External Examiner

Date...10...April...81.....

University of Alberta

To Jon

University of Alberta

ABSTRACT

A framework for modelling and simulation of phenomena in developmental genetics is presented in this thesis. Initially, formalisms for such modelling are introduced, using systems theory to provide a context for questions concerning model validity.

The biological and theoretical background of the project is discussed. Phenomena such as transdetermination, compartmentalization and positional information are described. Mechanisms for pattern formation such as chemical reaction and diffusion systems are investigated. Existing computer-oriented models are surveyed. The experimental frame for the project is discussed.

An underlying model for development is presented, and a modelling framework established, so that particular phenomena can be modelled by selecting options from a model pool. Structures are defined to enable representation of individual cells, groups of cells in structural relationships, and developmental systems. The modelling framework is multilevelled and of mixed type.

The implementation of the modelling framework in ALGOLW is described. The validity of approximations necessary in representing the modelling framework in iterative form is discussed. Such approximations include both discretization, of the time base and continuous state variables, and replacement of parallel by sequential processing.

University of Alberta

Particular experiments performed in developing modelling options are described. The control of growth is investigated, and options provided for locally mediated cell division. Intercellular communication by means of reaction and diffusion is modelled by means of numerical approximation, and it is shown that the concentration patterns predicted theoretically are observed at critical sizes of cell structures. The option of modelling the sequence of patterns by means of analytic solutions is provided for rectangular structures. A model is given for determination, wherein a cell's genetic information is affected on recognition of a critical value in its local conditions. Finally, models for segmentation in the blastoderm of Drosophila and for the development of the chick wing are presented, using options within the modelling framework.

Contributions to developmental genetics and to the art of modelling and simulation are discussed. Finally, the formal context employed in developing the modelling framework is evaluated in terms of its success in providing a mathematical environment for the study of developmental genetics.

University of Alberta

University of Alberta

ACKNOWLEDGEMENTS

I wish to thank Dr. Hugo Martinez, my external examiner, and Drs. John Tartar, Michael Russell, and Kenneth Morgan, the other members of my examining committee. Their interest and encouragement has been greatly appreciated.

I owe a deep debt of gratitude to my supervisor, Dr. Jeffrey R. Sampson. His guidance and continuing support have been invaluable.

I have been helped and encouraged by my family and friends, and in particular by my husband Jon, who typed most of this thesis.

This work has been supported in part by a National Research Council of Canada Postgraduate Scholarship and by a University of Alberta Dissertation Fellowship.

TABLE OF CONTENTS

Chapter Page

1. Introduction 1

1.1 Motivation 1
1.1.1 The Role of Mathematics in Genetics ..2
1.1.2 Scientific Modelling and Simulation ..7
1.1.3 Relations to Computing Science 8
1.2 Organisation of the Thesis 10
1.3 Summary of Goals 12

2. Modelling and Simulation 13

2.1 The Five Elements 13
2.1.1 The Real System 14
2.1.2 The Experimental Frame 14
2.1.3 The Base Model 15
2.1.4 The Lumped Model 15
2.1.5 The Computer 15

2.2 Formal System Specification 17
2.2.1 Time Base 17
2.2.2 State Set 17
2.2.3 Input and Output Sets 18
2.2.4 System Behavior 18

2.3 Manipulation of the Time Base 19

2.4 Heirarchy of System Specifications 22
2.4.1 I/O Relation Observation 22
2.4.2 I/O Function Observation 23
2.4.3 Input Output System 23
2.4.4 Iterative Specification 25
2.4.5 Structural Descriptions 26

2.5 Model Validity 26
2.5.1 I/O Relation Observation Morphism 27
2.5.2 I/O Function Observation Morphism 28
2.5.3 I/O System Morphism 29
2.5.4 Iterative Specification Morphism 30
2.5.5 Structured Specification Morphism 30

2.6 The Twelve Postulates 31

2.7 Categories of Models 34
2.7.1 Discrete Time Models 35
2.7.2 Discrete Event Models 35
2.7.3 Differential Equation System Specification 37

2.8 Simulation Techniques and Facilities 38
2.8.1 Continuous Time Simulation 39

University of Alberta

TABLE OF CONTENTS

Chapter Page

2.8.2 Discrete Event Simulation 40

2.9 Summary 42

3. The Real System 44

3.1 Genetic Information and Development 45

3.2 Disc Studies in Drosophila 50
3.2.1 Transdetermination and Codes 50
3.2.2 Positional Information and Gradients .53
3.2.3 Compartmentalization 56

3.3 The Mechanics of Pattern Formation 60
3.3.1 Active Transport 62
3.3.2 Phase Gradients 62
3.3.3 Cellular Contacts 63
3.3.4 Cell Sorting 63
3.3.5 Reaction-Diffusion Kinetics 64
3.3.6 Applications of Reaction-Diffusion Kinetics 71
3.4 Computer Oriented Models 78
3.4.1 Models Based on L-Systems 78
3.4.2 Growth Models in Array of Cells 80
3.4.3 Intracellular Growth Models 83

3.5 The Experimental Frame 83

4. The Modelling Framework 86

4.1 The Underlying Base Model 87

4.2 The Common Elements of Lumped Models 95
4.2.1 Characterization of Individual Cells .96
4.2.2 Characterization of a Group of Cells with the Same Genetic Code 97
4.2.3 Characterization of a Developmental System 98
4.2.4 Cell Division 98

4.3 Formal Characterization of Lumped Models 99
4.3.1 Specification of Individual Cells 99
4.3.2 Specification of Code Groups 100
4.3.3 Specification of a Developmental System 101

4.4 Variable Elements for Individual Models 102

University of Alberta

TABLE OF CONTENTS

Chapter Page

5. Implementation 103

5.1 Choice of Programming Environment 103

5.2 Treatment of the Time Base 106

5.3 Data Structures 106
5.3.1 Cells 106
5.3.2 Groups of Identical Genetic Nature ...111
5.3.3 The Developmental System 112

5.4 Sequential and Parallel Processing 112

5.5 Representation of Growth 115

5.6 Problems of Economization 116

5.7 Validity of the Implementation 119

6. Experiments and Results 120

6.1 Experimental Sequence 120

6.2 Modelling Growth 121
6.2.1 Formal Specification of the Growth Model 122
6.2.2 Sequential and Parallel Processing ...128
6.2.3 Growth of Competing Structures 138
6.2.4 Modelling the Division Decision 139
6.2.5 Modelling the Division Direction 150
6.2.6 Summary of Experimentation with Growth 151

6.3 Modelling Intercellular Communication 152
6.3.1 Formal Specification of the Communication Model 157
6.3.2 Experiments with Numerical Solutions .161
6.3.3 Modelling Using Analytic Solutions ...175

6.4 Modelling Growth, Communication and Determination 183
6.4.1 Formal Specification of a Determination Model 185
6.4.2 Growth and Communication 187

6.5 Modelling Developmental Systems 194
6.5.1 Modelling the Blastoderm of Drosophila 195
6.5.2 Modelling the Chick Wing 204

University of Alberta

TABLE OF CONTENTS

Chapter Page

6.6 Conclusion 220

7. Conclusions 221

REFERENCES 230

University of Alberta

LIST OF FIGURES

Figures Page

3.1 Kauffman's assignment of bistable switches 52
3.2 Bryant's gradient model 55
4.1 Coding in subdividing populations 88
5.1 State record for a cell 108
5.2 Position record for a cell 108
5.3 Neighbourhood relationships between cells 110
5.4 Shape record list for a code group 113
5.5 Group record list for a developmental system 113
5.6 Division of a cell 117
6.1 Growth with randomized processing list 133
6.2 Growth with unsorted processing list 134
6.3 Growth with geometrically ordered processing list 135
6.4 Growth from irregularity to convexity 137
6.5 Growth near an impenitrable barrier 140
6.6 Growth along and around a barrier 141
6.7 Successive time steps in clocked growth 143
6.8 Division in response to the value of a local function.....144
6.9a Some cells with odd clock settings will divide ..147
6.9b 10 growth cycles from 7 initial cells 148
6.10 Lineage 7 divides without control 153
6.11a First cosine pattern arises in one dimensional, two species system 166
6.11b First cosine pattern established, species one ..167
6.11c First cosine pattern, species two 168
6.12a Second cosine pattern, species one 169
6.12b Second cosine pattern, species two 170
6.13 Third cosine pattern, species one 171
6.14 Fourth cosine pattern, species one 172
6.15a Analytic solution, x direction (pattern one) ...177
6.15b Analytic solution, z direction (noncritical) ...178
6.15c Nodal pattern 179
6.16a Analytic solution, x direction (pattern one) ...180
6.16b Analytic solution, z direction (pattern one) ...181
6.16c Nodal pattern 182
6.17a Growth through cosine patterns, initial configuration 188
6.17b With 10 cells 189
6.17c With 14 cells 190
6.17d With 20 cells 191
6.17e With 40 cells 192
6.18a Segmentation in the blastoderm, nodal pattern one 198
6.18b Assignment of code 199
6.18c Nodal pattern two 200
6.18d Assignment of code 201
6.18e Nodal pattern three.....202
6.18f Assignment of code 203
6.19 Elliptic cross-sections in the wing 206
6.20a Limb bud, nodal pattern, phase one 210

University of Alberta

LIST OF FIGURES

Figures Page

6.20b Limb bud, cell determination, phase one 211
6.20c Growth, phase one 212
6.20d Nodal pattern, phase one 213
6.20e Cell determination, phase one 214
6.20f Growth, phase one 215
6.20g Nodal pattern, phase one 216
6.20h Cell determination, phase one 217
6.21 Cell determination, phase two 218
6.22 Cell determination, phase three 219

University of Alberta

CHAPTER ONE

Introduction

University of Alberta

1.1 Motivation

Developmental genetics is a vital area in current biological research. The genetic code has been deciphered, and investigations proceed into the means by which this inherited information stored in the cell is translated into biochemical activity. In a developing organism, this translation of genetic information results in a sequence of changes in cells and their progeny leading to a cooperative population of many cell types, that is, a viable organism. The genetic material in a cell may be regarded as a specification of all chemical activities possible in this organism: during development, each cell has operative only parts of this material. Which part of the genetic material is operative at a given time depends on the operational history of the cell (its initial state and all succeeding

states), any external influences, and its position and relationship to other cells in the organism. An immediate comparison may be made to the operation of an algorithm in a computer: which portion of the algorithm is in action depends on operation of the algorithm to date and inputs to which the computer may be connected.

The analogy between developing organism and computer is developed into a simulation project in this thesis. The project is a response to certain needs both in genetics and in that area of computing science known as modelling and simulation.

University of Alberta

1.1.1 The Role of Mathematics in Genetics

In the physical sciences mathematics has played an essential role in providing a formal and exact context for the formulation of theories and their experimental verification. The aim of biological science is similar to that of the physical sciences, that is, the development of generalized, unifying theories about its own domain. While much of modern mathematics has developed in interaction with the physical sciences, many important milestones in biology, and in particular in genetics, have been achieved when conventional mathematics has been applied in biology.

Scientific biology may be considered to have begun when it was recognized that the metabolic functions and transformations in organisms are reducible to simple chemistry. The modern biochemistry of enzymatic reactions,

based on the work of Haldane (1930), consists of quantitative reaction laws based on the mathematics of physical chemistry. Differential equations express changes of concentration of interdependent chemical species. Thus, at the molecular level, biology has a useful mathematical formalism.

The mathematical theory of evolution and natural selection, in which the pioneers were Wright (1931), Haldane (1932) and Fisher (1930), rests on the recognition of the gene as a heritable unit and the application of stochastic mathematics to genetic populations. Thus at the level of the genetic population there is a useful mathematical formalism.

Consideration of possible mechanisms by which genetic information may be interpreted differentially led to the discovery of phenomena such as end product inhibition, in which the end product of a synthetic reaction combines with a catalyst of the same reaction to modify the reaction rate. Thus control systems theory began to be used in the context of genetic mechanisms, and in fact attempts have been made to describe the whole organism as an adaptive control system (Reiner, 1968; Rosen, 1972). Thus both at the molecular and at the organic level, systems theory provides a useful mathematical formalism. The use of this formalism will be discussed in greater detail in Chapter Two.

Dynamic processes by which form and pattern are generated in biological systems may be grouped under the

University of Alberta

name morphogenesis. The fundamental question in such processes is how patterns arise autonomously in initially uniform situations. A theory first developed by Rashevsky (1960) and Turing (1952) suggests that the uniform configuration must be the result of an unstable equilibrium which may be transformed by random fluctuations into a nonuniform stable state. Mathematical formalisms which attempt to deal with such phenomena include bifurcation theory (Nicolis and Prigogine, 1977), catastrophe theory (Thom, 1972), and stability theory (Gmitro and Scriven, 1966). These theories will be discussed in greater detail in Chapter Three. In developmental genetics the scope of such a treatment will usually be the morphogenetic field, which is of the order of 100 cells (Wolpert, 1978; Crick, 1970).

There are many examples of attempts to apply quantitative methods to individual phenomena in developmental genetics; some of these will be discussed in Chapter Three.

It appears likely, therefore, that any single phenomenon in genetics may be explicable in terms familiar to the mathematics of physical science. At the molecular level there is enzyme kinetics; at the level of the cell population there is statistics; at the level of the morphogenetic field there may be a mathematical formalism for equilibrium states. However development is a phenomenon occurring at all these levels, and it has an additional

University of Alberta

characteristic. It is a process occurring in and affecting a structure of cells.

Essential features of pattern formation in a growing organism are the relative positions of cells and their intercommunication. The mathematics of classical mechanics deals with points, while that of modern physics deals with fields; neither context is sufficient to describe development.

The computing scientist may suggest that the "cellular space", wherein independently functioning "cells" are tied into a structure by means of connections to neighbours, is the mathematical tool suited to development. In addition there is an extant body of mathematics related to cellular automata; however there are features of a developing organism which make conventional cellular automata theory insufficient. Prominent among these are, firstly, nonuniformity of function in the cells of increasing complexity as cells differentiate, local functioning not being subject to local control, and secondly, the continual growth and change in the structure as cells divide or move. The structure of cells is dynamically subject to controls which are mediated at the level of the organism, or at least the morphogenetic field.

A mathematical context for development must be capable of mirroring the interaction between process (the dynamic aspect) and structure (the static aspect). Furthermore, it should be capable of invoking the mathematical tools

University of Alberta

mirroring events at molecular and collective levels. This thesis outlines the development of a project in which systems theory is used in alliance with a cellular space to attempt this task.

An individual biological cell can act as an automaton (for example, in determining and retaining a cell type marker) or as a location in a field (for example, in determining local concentration of a freely flowing chemical). Thus a developing organism has both continuous and discrete aspects (Lewis, 1976). The operation of field-like phenomena can "feed back" to change structure and local functioning in the cellular space. While the technique will not be discussed in greater detail here, two observations may be made. Firstly, a heirarchy of control is implied, that is, the interaction of process and structure occurs at a non-local or "higher" level. Secondly, it is likely that various options should be present in any situation, for example, different mathematical theories of cell division or of cell intercommunication may be desirable. A "pool" of techniques should therefore be provided.

A computer is an ideal tool in a situation where many cells, each capable of various behaviors, interact in a dynamic fashion. Its large storage capacity allows consideration of many factors, while its capacity to generate data dynamically following an iterative program specification allows for dynamic testing of theories of development.

University of Alberta

Developmental genetics, in summary, will benefit from the evolution of mathematical formalisms suitable to its particular features. When individual experiments can be related to such a context, more meaningful testing of theories of development will result. Furthermore, in a field where the number of parameters and options is vast, more fruitful direction for experimentation can be expected in a unifying context (Reiner, 1968). This thesis establishes a mathematical context for developmental experiments which includes computer simulation.

1.1.2 Scientific Modelling and Simulation

A hypothesis about a natural phenomenon may be expressed in mathematical (formal, symbolic, and exact) terms. How is this formulation related to, firstly, the phenomenon it is supposed to represent, and secondly, the refinements through which the experimenter may put it in testing the hypothesis?

A natural phenomenon is known to the experimenter solely through its behavior (measurements on parameters, response to stimuli or inputs, etc). He makes a hypothetical construct which, he hopes, represents reality in some fashion; usually, it is designed to generate similar behavior.

Modelling and simulation are terms used in describing the process wherein a hypothesis is translated into formal terms which may then be used in an iterative manner to

University of Alberta

generate artificial behavior (to be compared to the real data available to the experimenter). The process becomes a science when the relationship between model and reality, and between one modelling formalism and another, can be given in equally formal terms. Only thus can the intended scope of a model be defined and its significance properly assessed.

Ziegler (1976) has proposed a theory of modelling and simulation which suggests the manner in which a modelling experiment should be constructed and formalized if questions of validity are to be addressed directly. Modelling in developmental genetics, with its multiplicity of parameters and influences, appears an especially challenging task. This thesis will use Zeigler's methods wherever possible, both as an aid in a complex project and in an attempt to assess the applicability of the theories. Zeigler's theories are particularly suited to this project because they are couched in systems theoretic terms; thus the dynamic aspects of the project can be included naturally.

University of Alberta

1.1.3 Relations to Computing Science

Simulation is a technique used widely by computing scientists, whether in the context of efficient data processing (for example, queueing models) or in more applied fields (for example, information retrieval or ecological models). A project such as this one, though specialized in aim, will provide contributions to the general art of simulation itself.

The project is concerned with dynamically occurring changes in structure and local function in a cell space. In simulation and in other areas of computing science where the cell space formalism is used, considerable power and extension of applicability can be expected if methods of dealing with change and non-uniformity are developed.

Biological information processing systems provide a rich environment for both multilevelled and mixed-type modelling. In a developmental context modelling can be done at the various levels of the organism, part of the organism, the morphogenetic field, the cell, or the intracellular genetic apparatus, or at a combination of these levels. At different levels different modelling techniques may be appropriate; furthermore, it may be possible to model the same phenomenon at varying levels of "resolution", with different techniques, and thus have a means of checking results (Sampson, in preparation). As discussed above, a developmental system contains both continuous elements (for example, time and field-like influences) and discrete elements (for example, certain cellular characteristics), and thus must be modelled as a system of mixed type. Paradigms for multilevelled, mixed type modelling are essential for many real systems of interest, and therefore successes in these areas may be expected to contribute to the art of simulation.

Computing science is the science of successful information processing. Biological information processing

University of Alberta

has been proven successful through the process of evolution, and therefore may have ideas to contribute to artificial systems. In particular, biological systems store information compactly (the memory element being at the molecular level), are reliable for the lifespan of the system, and above all behave adaptively. Developmental systems exhibit particularly powerful adaptation on the basis of experience, as cells and structures develop specialized functions. Artificial information processing looks toward more compactness, greater reliability, and adaptibility in both hardware and software. While the mechanisms of biological systems may not be directly and currently applicable in artificial systems, an understanding of such mechanisms may be expected to contribute to the design of the artificial system of the future.

In summary, then, simulation in developmental genetics has special features and difficulties which, if overcome, may be expected to contribute both to the art of simulation and to computing science.

1.2 Organization of the Thesis

Since Zeigler's theory of modelling and simulation (Zeigler, 1976) is used as a formal background in developing this project, Chapter Two is devoted to a concise exposition of the theory. The theory divides such an enterprise into five elements: the real system, the experimental frame, the base model, the lumped model, and the computer. Chapters

University of Alberta

Three, Four and Five explore these elements in the context of the present project.

Chapter Three, entitled The Real System, introduces the necessary genetics and the actual areas in which the project operates. Since the state of knowledge and direction of experimentation in the subject seems inextricable from models under current consideration, the nature of existing models is also described in this chapter; these include both purely genetic modelling and work involving simulation. The chapter ends by discussing means by which experimental frames will be selected for particular modelling experiments.

Chapter Four describes the base model for developmental systems (the complete hypothetical model capable of generating all possible behaviors in all experimental frames). The modelling framework is described in informal and formal terms. Those elements of particular lumped models (simplified versions of the base model viewed in particular experimental frames) which are held in common are then described informally and formally.

In Chapter Five translation of the lumped models into iterative form is discussed. Such issues as the choice of programming environment, treatment of the time base, data structures, and problems of economization are discussed.

Chapter Six describes the implementation of particular lumped models (models of certain developing systems). The results of simulations are presented and discussed, with

University of Alberta

attention to questions of validity and general implications.

Chapter Seven evaluates the success with which the various aims of the project have been achieved.

1.3 Summary of Goals

The most immediate aim of the project is to provide a useful, functioning framework for testing models in developmental genetics. Achievement of this goal will depend on reaching many small individual goals in obtaining successful results in particular experiments. It is only in the context of such success that larger implications of the project may be discussed. These will include evaluation of the extent to which a satisfying mathematical context for development has been achieved, and demonstration of the role which the computer can play in such a context. In addition it will be possible to assess Zeigler's theory of modelling and simulation as a practical strategy in projects of this complexity. Further, contributions to the art of simulation and to computing in general are expected.

University of Alberta

CHAPTER TWO

Modelling and Simulation

University of Alberta

Three components of the process known as "modelling and simulation" are the real system, the model, and the computer. A model is constructed as a conjecture about the nature of a real system, and a simulation built to generate the behavior of the model. Zeigler (1976) has developed a framework for defining these components and establishing relationships between them in theoretical terms. The essence of his treatment is presented here, with the aim of introducing the present author's models and simulation experiments in perspective.

2.1 The Five Elements

Zeigler expands the three components to five conceptually distinct elements of modelling and simulation. These are the real system, the experimental frame, the base

model, the lumped model, and the computer.

2.1.1 The Real System

The real system is a natural or artificial source of behavioral data. The data are specified as values of descriptive variables, which may be observable (measurable by current techniques) or nonobservable (inaccessible to the investigator).

Observable variables are further classified as either input variables (those getting their values from influences external to the system) or output variables (those receiving values as the result of values of input variables). Over a given period of time, sequences of values of all observable variables may be collected into input and output segments for each variable. An I/O segment pair is the combination of an input segment and an output segment for a certain time interval. The set of all possible I/O segment pairs is all that can be directly known about the real system, and is known as the "I/O behavior" of the system.

2.1.2 The Experimental Frame

Complexity of the real system or limited resources available to the experimenter usually limit access to the real system to be studied. Associated with this limiting experimental frame is a subset of the I/O behavior of the real system.

University of Alberta

2.1.3 The Base Model

The base model is a set of instructions capable of generating all the I/O behavior of the real system. All variables in all possible experimental frames are accounted for. In other words, the base model is a hypothetical complete explanation of the real system.

2.1.4 The Lumped Model

Within an experimental frame, models may be constructed which are less complex than the base model and yet capable of generating the subset of I/O behavior of that frame. Various simplification procedures may be used to produce a lumped model from the base model; these include eliminating descriptive variables, approximating deterministic, causal sequences affected by many factored stochastic events, coarsening the range of descriptive variables, and pooling (or "lumping") components and their descriptive variables. The hazards of simplification and degrees of validity of the resulting lumped models will be discussed when a formal terminology for model description has been introduced.

2.1.5 The Computer

The computer is a device which can take a suitable specification of the lumped model and generate its I/O pairs. The iterative execution of the model instructions is known as simulation.

In terms of these five elements, "modelling" is the

University of Alberta

process of constructing a lumped model which is valid with respect to the real system in the context of an experimental frame. "Simulation" is the process of generating the behavior of a model by means of the execution of a computer program.

The computer program itself can be regarded as a specification of the lumped model. However, such specification is machine and language dependent, and its structure may reflect computational facility rather than the inherent nature of the model. It is desirable to obtain a formal model description which is directly implementable and also provides a means of comparison between models (for example, when simplification procedures are being employed to produce more lumped models). Furthermore, if any model can be expressed in a standard form, its salient distinguishing features can be immediately recognized and decisions regarding means of simulation made most efficiently. Zeigler describes a hierarchy of model specifications of increasing rigour, from verbal descriptions of components with their descriptive variables and interactions, to "iterative system specifications". The latter are system theoretic specifications and amount to description of a model in a certain canonical form. The construction of such a specification is described below.

University of Alberta

2.2 Formal System Specification

A dynamic system is a set of components with attributes which vary with time and interact with each other. The following fundamental elements define a system.

2.2.1 Time Base

A time base is a set T which is isomorphic either to the reals, in which case T is said to be continuous, or to the integers, in which case T is said to be discrete. In both continuous and discrete form, T respectively obeys the following properties of the reals and the integers.

Linear ordering allows definition of past, present and future. The stipulation that T be an Abelian group under addition allows consideration of translation of events in time. The requirement that T be unbounded above and below allows consideration of arbitrarily large time segments. Chronology of events is preserved (under translation) if addition is required to be order preserving ($t1 < t2 \rightarrow t1 + t3 < t2 + t3$ for all $t3 \in T$).

University of Alberta

2.2.2 State Set

Associated with the components of the system are descriptive variables which fully define them. At any point in time, the state of the system is given by the values of these variables. The state variable set, Q, is the minimal set which specifies the system as it undergoes changes. The values of these variables, together with the inputs, at any

time t determine (uniquely in the case of models with no stochastic elements) the values of all descriptive variables at all future times.

2.2.3 Input and Output Sets

A nonautonomous system exists in some environment, with which it interacts via inputs and outputs. The input set, X, is a collection of variables which influence, but are not influenced by, the system. The output set, Y, is some subset of the descriptive variables of particular interest in the given system.

The state, input and output variables may be finite, discrete, continuous or mixed, and the time base may be discrete or continuous. Thus a system may be categorized by the nature of its time base and variables.

University of Alberta

2.2.4 System Behavior

A system in some state at a given time experiences a set of input values. It undergoes a transformation to a new state and produces certain output values. Such behavior can be described in terms of two functions, the transition function, f, and the output function, g:

$$f: Q \times X \rightarrow Q$$

$$g: Q \rightarrow Y$$

The functions f and g may be deterministic, if each element in Q x X or Q maps to only one element of Q or Y, or nondeterministic, if some domain elements each map to more

than one range element. If the choice of range element is determined stochastically, then the system is termed stochastic.

A complete system description, then, is given by the structure <T,X,Q,Y,f,g>. When some of the elements are not known, or not known completely, a less detailed structure must be used to describe the system.

A "model" is any system which corresponds in some way to another system. The modeller's intention is to develop a fully specified system as a model of one which is not fully understood and therefore not fully specified. For purposes of simulation, the complete specification of a model should be in iterative form, that is, capable of step-by-step computation from an initial specified state. A hierarchy of system specifications can be described in which an increasing amount is known or specified, culminating in iterative specification.

University of Alberta

2.3 Manipulation of the Time Base

Inherent in the notion of iterative specification is the idea of manipulations on the time base. At each computational instant, inputs are treated as arriving now (t = 0), and state transitions are immediate. Furthermore, only the behavior relevant to the current time-step length is to be considered. As a preliminary to the description of specification hierarchies, operations on and assumptions about the time base will be presented.

Let Z be one of X, Q and Y. Let <t1,t2> be a time interval. Here < > implies ( ), ( ],[ ), or [ ]. Then let w be a mapping from <t1,t2> to Z, that is, an input or output segment or a state trajectory from t1 to t2.

$$w:\langle t1,t2\rangle \dashrightarrow Z$$

w is known as a segment in (Z,T). If w1, w2 are segments in (Z,T) and w1: <t0,t1> --> Z and w2: <t1,t2> --> Z they are said to be contiguous. A composition operator (•) may be defined for contiguous segments:

$$w1 \bullet w2 = \begin{cases} w1 \text{ for } t \in \langle t0,t1\rangle \\ w2 \text{ for } t \in \langle t1,t2\rangle \end{cases}$$

If a unique value is assigned at t1 (according to whether > means) or], and < means (or [) then composition is associative: let w1 and w2 be contiguous, and also w2 and w3. Then

$$(w1 \bullet w2) \bullet w3 = w1 \bullet (w2 \bullet w3)$$

The notion of composition allows formalisation of the idea of successive input experiments. A segmentation operator may also be defined, allowing experiments to be decomposed into smaller input setments. In addition, a translation operator allows a segment to be applied at a different time.

Given a set of segments W, a subset of (Z,T), it is desired to find a set of segments G such that compositions of segments in G yield the segments of W. Such a set G is called a set of generators for W, or G is said to generate W. The set {w1,w2,...,wn} is a decomposition of w by G if

University of Alberta

wi $\in$ G, i=1,2,...,n, and w = wi•w2•...wn.

In the interests of iterative representations it is desirable to obtain unique (canonical) decompositions of segments. Such decompositions may be obtained by a process called maximal length segmentation (mls).

The process of finding an mls decomposition for w $\in$ W may be understood more clearly as follows.

(1) Find wi, the longest generator in G which is also a left segment (a segment mapping from the same initial time as w) of w.

(2) Repeat for what remains of w after the left segment wi has been removed. If the process stops after n repetitions, then w = wi•w2•...wn and {wi,...wn} is the mls decomposition.

It is not true that mls decompositions always exist (that is, that the process outlined terminates). Zeigler presents a theorem for conditions guaranteeing mls decompositions.

If G generates W and for each w $\in$ W an mls decomposition by G exists, G is said to be an admissible set of generators for W.

The above discussion of operations on segments accomplishes the ability to move, combine, and decompose them canonically under stated conditions. This framework may be used in describing the hierarchy of system specifications.

University of Alberta

2.4 Hierarchy of System Specifications

There is a heirarchy of specifications for systems of increasing power. The more is known about a system, the more detailed and definite is the description that can be specified. At the level where a system is known only by its behavior, an I/O relation observation can be formulated.

2.4.1 I/O Relation Observation

At the most basic experimental level, a system is known by its input/output pairs. An I/O relation observation is a structure <T,X,W,Y,R> where

T is a time base

X is an input value set

Y is an output value set

W is an input segment set, a subset of (X,T)

R is a relation, a subset of W x (Y,T)

where (w,p) $\in$ R => domain(w) = domain(p). (that is, w and p are input and output segments over the same observation interval).

R is the collection of I/O pairs, and the structure is a purely behavioral description. Nothing is said about internal state of the system or about means of generating output.

University of Alberta

2.4.2 I/O Function Observation

The set R may be partitioned if different initial states of the system are related to different sets of I/O pairs. An I/O function observation is a structure <T,X,W,Y,F> where T,X,W and Y are as before and F is a set such that

f ∈ F ==> f is a subset of Wx(Y,T) and f is a function. (If p = f(w), then domain(p) = domain(w)).

Each f refers to a distinct initial state. The collection of all I/O pairs generated by all f's is the set R. An I/O relation observation may be obtained from an I/O function observation, then, but the reverse is not true as information about initial states is needed for an I/O function observation.

University of Alberta

2.4.3 Input Output System

An I/O system is a structure

S = <T,X,W,Q,Y,f,g>

where

T is a time base

X is the input value set

W is a subset of (X,T), the input segment set.

Q is the state set

Y is the output value set

f is the state transition function

g is the output function

subject to the constraints

(1) W is closed under composition, that is, for all contiguous segments $w1, w2 \in W$, $w1 \bullet w2 \in W$.

(2) f and g are deterministic mappings

$$f: Q \times W \rightarrow Q$$

$$g: Q \rightarrow Y$$

(3) f obeys the composition property, that is, for every contiguous pair of segments $w1, w2 \in W$

$$f(q, w1 \bullet w2) = f(f(q, w1), w2)$$

The significance of constraint (1) is the necessity for compound I/O experiments to be allowed in a system. Constraint (2) embodies the requirement of a deterministic system that, given an initial state and a particular input segment, the system will inevitably undergo transition to the same final state. Constraint (3) states that the fragmentation of an input segment does not affect the final state reached through operation of the transition function throughout the segment. The constraints formalize the ideas intuitively necessary for a deterministic dynamic system, those of re-initialization, repetition, interruption and re-starting of an experiment.

Given the structure S it is possible to generate the behavior of a system. That is, it is possible to obtain an I/O function observation and an I/O relation observation which are complete in the sense of containing all possibly observable I/O pairs.

University of Alberta

2.4.4 Iterative Specification

An iterative specification of a system is a structure

G = <T,X,W,Q,Y,f,g>

where T, X, Q, Y, and g are as before, and W is a subset of (X,T), the input segment generator set, and f: Q x W --> Q is the single segment transition function, subject to

(1) W is an admissible set of generators.

(2) f has the composition property.

The iterative specification structure gives only the set generating the input segment set for the corresponding system, and the transition function for each segment in the generating set. Thus if w has an mls decomposition by W of w1•w2•w3•...•wn, then the state of the system after w is given by applying the single segment transition function to each wi in turn, i.e. if initial state is q(0), then let

$$q(1) = f(q(0),w1)$$

$$q(2) = f(q(1),w2) \quad \text{etc.}$$

and the final state reached is q(n) = f(q(n-1),wn)

The utility of an iterative specification is twofold. First, it is a minimal description since the single segment transition function operates on segments which may occur in the decomposition of many input segments. Second, a transition function operating iteratively to achieve a compound effect is useful in bridging the gap between model description and simulation program.

A time-invariant I/O system specification can be obtained from an iterative specification by compounding the

University of Alberta

single segment transition function. An iterative specification contains more information in that it is known that the input segment set can be decomposed and the transition function can be represented in terms of action on single segments.

2.4.5 Structural Descriptions

Zeigler gives a means of describing systems which can be specified as interacting networks of component sub-systems. The descriptions consist of specifications of the component sub-systems together with an expression labelling the components and, for each one, showing those others which it affects, that is whose input segments it generates. Such structural specifications correspond most closely to informal model descriptions in terms of interacting components.

Translation of a specification at the highest level into one at the lowest is the formal equivalent of computer simulation of a model. Step-by-step successive states and outputs are obtained using an iterative description, and thus behavioral data is collected.

2.5. Model Validity

A "model" may be thought of as one system which attempts to characterise some of the features of another system. Successful modelling depends on the establishment of relationships between such systems. The validity of a base

University of Alberta

model in representing the real system, of a lumped model in representing the base model, and of the simulation program in implementing the lumped model must all be investigated via such relationships.

At each of the levels of specification of systems a "preservation relation" or "morphism" may be stated which establishes a correspondence between features of two systems described at that level. To the hierarchy of specification levels corresponds a hierarchy of morphisms. This hierarchy is strict in that lower level morphisms do not imply higher level ones. For example, systems which display the same kind of input-output behavior may generate this behavior in entirely different manners.

University of Alberta

At a particular level of specification, morphisms may be described between systems S1 and S2. Let S1 be larger than S2 in the sense that S1 is capable of doing anything that S2 can, but not vice versa. A mapping h from S1 to S2 is called a "surjective" or "onto" map, and ensures that all the behavior of S2 is being covered by S1. A mapping from S2 to S1 is called an "injective" map, and embeds S2 in S1, or in other words, isolates the appropriate part of S1 to relate to S2.

2.5.1. I/O Relation Observation Morphism

Let S1 be given by <T1,X1,W1,Y1,R1>. Let S2 be given by <T2,X2,W2,Y2,R2>. Let h: W2 --> W1. h is injective and can be referred to as an encoder. If w2 $\in$ W2 is applied to S2,

then h(w2) = w1 ∈ W1 tells which input segment to apply to S1.

Let k: (Y1,T1) --> (Y2,T2). k is surjective and can be referred to as a decoder. An output segment p1 ∈ (Y1,T1) is known to represent k(p1) = p(2) ∈ (Y2,T2).

The notion that the I/O behavior R2 of S2 can be found in the I/O behavior R1 of S2 can be formalised as follows:

An "I/O Relation Observation Morphism" from S1 to S2 is a pair (h,k) such that

1. h: W2 --> W1
2. k: (Y1,T1) --> (Y2,T2) and is onto.
3. for every (w2,p2) ∈ R2 there exists a (w1,p1) ∈ R1 such that w1 = h(w2) and k(p1) = p2.

Thus every I/O segment pair of S2 has its corresponding pair for S1 and the mapping of one pair to the other is established. If input w2 is applied to S2, the behavior of S2 may be studied by observing the behavior of S1 on application of h(w2), and decoding the result.

2.5.2 I/O Function Observation Morphism

An I/O function observation morphism from S1 = <T1,X1,W1,Y1,F1> to S2 = <T2,X2,W2,Y2,F2> is a pair (h,k) where

1. h: W2 --> W1
2. k: (Y1,T1) --> (Y2,T2)
3. For each f2 ∈ F2 there is an f1 ∈ F1 such that f2 = k•g; that is, for all w2 ∈ W2, f2(w2) = k(f1(h(w2))).

University of Alberta

The effect of applying w2 to S2 may be studied by coding it to obtain w1 = h(w2), applying this to S1 and obtaining p1 = f1(w1) = f1(h(w2))), and decoding to p2 =v k(p1) = k(f1(h(w2))).

2.5.3 I/O System Morphism

A system morphism from a system S1 = <T1,X1,W1,Q1,Y1,f1,g1> to S2 = <T2,X2,W2,Q2,Y2,f2,g2> is a triple (h,j,k) where

1. h: W1 --> W2
2. j: Q --> Q2 where Q is a subset of Q1, and j is onto.
3. k: Q1 --> Y2 and is onto.

and for all q ∈ Q, w2 ∈ W2,

4. j(f1(q,h(w2))) = f2(j(q),w2) (transition function preservation).
5. k(g1(q)) = g2(j(q)) (output function preservation).

Under a system morphism, there is not only behavior preservation but also there are structural relationships between the state spaces and the means by which state changes and behavior are generated.

If T1 = T2, X1 = X2, W1 = W2, Y1 = Y2 then S1 and S2 are called "compatible". Compatible systems S1 and S2 are homomorphic images of each other if (h,j,k) is a system morphism, h and k are identity maps, and Q = Q1; j is said to be a homomorphism. If j is one-to-one as well as onto it is an isomorphism. In essence, two systems are homomorphic if their state sets map completely into each other and if

University of Alberta

the action "encoding (S1 to S2) followed by transition (in S2)" results in the same state as the action "transition (in S1) followed by encoding (S1 to S2)".

Systems obeying a system morphism can be shown to obey behavioral morphisms as well.

2.5.4 Iterative Specification Morphism

A "specification morphism" is a triple (h,j,k) defined similarly to a system morphism, with the differences that h maps S2's generator set into S1, and that transition function preservation need only be checked for the generator set of the smaller system. It may be shown, by expanding each iterative specification into its system specifications through compounding the single segment transition functions, that a specification morphism implies a system morphism. Thus one of the major advantages of iterative specifications is the ability to check preservation relationships between them without expansion to system form and exhaustive checking of state transitions for all members of the input segment set.

2.5.5 Structured Specification Morphisms

A structured specification is one in which labels and interconnections are given to subsystems of a large system. Structural morphisms of two kinds are discussed by Ziegler (p. 276). A "weak structure morphism" is a homomorphism between the state structures of two structured systems, but

University of Alberta

while the global transition functions of the systems are related, local transition functions of the subsystems need not be. (Typically subsystems of the larger system S1 are aggregated into blocks, and each block is a subsystem of S2). In order to obtain a "strong structure morphism" between systems, it is necessary to ensure that local interaction of component subsystems is preserved.

The morphisms thus constructed play a fundamental part both in consideration of model (and simulation) validity and in facilitation of model construction. The next section presents Ziegler's twelve fundamental postulates (Zeigler, 1976) for successful modelling and simulation, and these postulates illustrate the utility of morphisms in constructing valid models. Ziegler discusses using morphisms to simplify the modelling task in later chapters of his book; as an example, large systems may be decomposed into manageable subsystems provided structural morphisms are maintained.

High level morphisms imply lower level ones; the significance of this is that if systems obey a high level morphism, it is known that they are behaviorally related without the need for generating behavior and comparing it.

2.6 The Twelve Postulates

Zeigler devotes a chapter to the statement and discussion of his twelve fundamental postulates (Zeigler, 1976, Chapter 11). The postulates are paraphrased below.

University of Alberta

Postulate 1: There exists a real system R which is identified as a set of potentially acquirable data.

Postulate 2: There exists a base model B that structurally characterizes this set of potentially acquirable data. B is a time-invariant transition system (system without output)

$$B = \langle T, X, (X,T), Q, f \rangle$$

Postulate 3: There exists a set of experimental frames {E} restricting experimental access to R.

Postulate 4: An experimental frame E is a structure $E = \langle W, Y, g, V \rangle$. That is, a subset W of (X,T) of input segments applicable in E, a set Y of output values observable in E, an output function g which maps base model states to outputs, and a set V of valid output (determined by experimental conditions of E).

Postulate 5: The real system observed within the experimental frame is structurally characterized by the base model in the frame, denoted B/E.

Postulate 6: $B/E = \langle T, X, W, Q, Y, f, g \rangle$ where

T is the time base of the base model

X is the input value set

W is the set of input segments allowed under E

Q is the set of base model states which lie within valid

University of Alberta

range for E together with 0, a symbol for states outside E.
Y is the valid output set together with 0, to which nonvalid outputs are collapsed.
f:Q x W --> Q is the transition function of the base model where it results in valid states for E; all other states obtained are collapsed to 0.
g:Q --> Y is the output function of E unless applied to a state not valid for E, in which case it maps to 0.

Postulate 7: The data potentially acquirable by observations of the real system within experimental frame E are identified with the I/O relation of B/E.

Postulate 8: The real system R is the set of data potentially acquirable by observation of all experimental frames.

Postulate 9: A lumped model is an iterative system specification M.

Postulate 10: A lumped model is valid for the real system in experimental frame E if its I/O relation is equivalent to that of B/E.

Postulate 11: A computer is an iterative system specification C.

University of Alberta

Postulate 12: A computer C is a valid simulator of a lumped model M if there is a specification morphism from C to M.

Postulates 10 and 12 define what is meant by "valid model" and "valid simulation", and the limited significance such assessments are meant to have. A lumped model valid for R in E is valid in terms of behavior; its own known structure may provide insight into structural properties of the base model, but such base model structure cannot be established empirically. However, there may be conditions under which structural inference about B may be made (Ziegler, 1976, Chapter 15).

2.7 Categories of Models

Systems may be categorized according to whether the time base is discrete or continuous, and variable sets are finite, discrete or continuous (or mixed).

Models in common use most often fall into three categories; discrete/discrete (sequential machines), continuous/discrete ("discrete event" models) and continuous/continuous ("differential equation" models). Simulation packages for models in the latter two categories are widely available.

University of Alberta

2.7.1 Discrete Time Models

A discrete time system specification, or sequential machine, is a structure

$$M = \langle XM, QM, YM, f, g \rangle,$$

XM, QM, YM subsets of the integers. Corresponding to M is an iterative specification G given by

$$G = \langle T, XM, W, QM, YM, fi, gi \rangle$$

where

$$T = I$$

$$W = \{w \mid w:(0,1) \dashrightarrow X\}$$

$$fi: Q \times W \dashrightarrow Q,\ fi(q,w) = f(q,w(0))$$

$$gi = g$$

Each $w: (0,1) \dashrightarrow X$ is uniquely defined by its value $w(0)$ because T is a discrete time base. Thus the generators in W associated with the iterative specification are in a one-to-one correspondence with the input values X. Thus the transition of the system under f in response to a sequence of input values from XM is equivalent to the sequence of transitions fi each responding to a single input from XM.

2.7.2 Discrete Event Models

A discrete event model is one in which discrete state changes occur at arbitrary time intervals. Events are predictable on the basis of previous events and the system's present state, and therefore can be scheduled.

A discrete event system specification is a structure

$$M = \langle XM, SM, YM, f, g, t \rangle$$

University of Alberta

where XM, YM are as before, and
SM is the set of sequential states
f is the "quasi"-transition function
g is the output function
t is the time advance function H
such that

1) $t:s \rightarrow Rt$ and $t(s)$ is the time the system is to remain in state s.

2) Let $QM = \{(s,e) \mid s \in SM, 0 \leq e \leq t(s)\}$
Let z be a special symbol not in XM. Then $f: QM \times (X \cup \{z\}) \rightarrow SM$ such that $f(s,e,z) = fz(s)$ for all $(s,e) \in QM$, where $fz:SM \rightarrow SM$.

3) $g:QM \rightarrow YM$

Change in the system is caused either by completion of time $t(s)$ and arrival of the next (scheduled) event, or by the arrival of an (unscheduled) external event. If no external event occurs in $t(s)$, symbolized by z, then transition is to the next state $fs(x)$. If an external event x arrives after state s has endured for time e, and $e \leq t(s)$, then a new state $f(s,e,x)$ is immediately initiated. This new state depends on e, and thus the full state set QM includes e.

An iterative form of the system specification is constructed by considering the generating segments and associated transitions. Let Wz be the set of segments of finite length where no external event occurs and Wx be the set of segments of finite nonzero length where no external

University of Alberta

event occurs except at zero. Then the generator set Wg is the union of Wz and Wx. The transition function can then be constructed by considering actions on the generating segments. Inclusion of the variable e in the state set, to show elapsed time in state s, means that in the iterative form the time advance function can be dropped and a continuous time base established. Simulation of such an iterative form is then straightforward.

2.7.3 Differential Equation System Specification

Many models of natural phenomena are specified in terms of instantaneous rates of change in the state variables: that is, state changes are not prescribed directly but must be computed as satisfying differential equations.

A differential equation specified system is a structure

D = <X,Q,Y,fp,g>

where

X is the input value set

Q is the state set

Y is the output value set

fp is the rate of change function

g is the output function

subject to

(a) X, Q and Y are real vector spaces

(b) deterministic responses

fp: Q x X --> Q

g: Q --> Y

University of Alberta

In order to convert this specification into iterative form it is necessary to assume a continuous time base T, and as input segment generators, the bounded continuous functions from finite intervals of T into X. Then fp can be interpreted in a manner specifying the single segment transition function.

Let w: <0,t1> --> X and q ∈ Q be a given input segment and a state. Then a segment P: <0,t1> --> Q is a solution associated with w and q if P(0) = q and dP(t)/dt = fp(P(t),w(t)) for each t ∈ <0,t1>. If exactly one solution exists for each pair (w,g), the specification is that of a deterministic system whose state trajectories are given by the solutions. (Conditions on fp which ensure the existence and uniqueness of solutions are known).

The transition function in the iterative specification of a unique solution differential equation system is given by the mapping carried out by the solution taking a given initial state into the corresponding final state at the end of the observation interval.

2.8 Simulation Techniques and Facilities

The modeller must take a model specification in iterative form and translate it into a program interpretable by the computer. If the model falls into the discrete event or differential equation category, the task may be simplified by the availability of simulation packages. The modules of such packages may have specification morphisms

University of Alberta

known to connect them to the model elements they represent. If the model does not fall into one of these categories, the programming task is increased unless the modeller happens to have access to locally available, previously written programs applicable to his work. Programming in a general purpose language may, however, result in a more intimate control over the process of simulation.

2.8.1 Continuous Time Simulation

Implementation of continuous models on a digital computer necessitates discretization of the continuous variables. In the case of differential equation models, discretization of the time base implies approximation of differential equations by difference equations. The experience of the numerical analyst in choosing approximation techniques for a given differential equation may then be applied to maintaining a specification morphism between the transition function in continuous time, as given by the solution of a differential equation, and the transition function in discrete time, as given by the solution of the approximating difference equation by numerical integration techniques.

Numerical integration packages, which contain programmed versions of integration methods, and associated documentation assisting the modeller in choosing a method suited to his needs, are widely available. Techniques for estimating errors of approximation are known for such

University of Alberta

methods. Furthermore, simulation languages for continuous modelling such as CSMP and DYNAMO are available which perform the continuous time to discrete time conversion for the modeller. Successful use of such languages depends on the appropriateness of the approximations chosen for the particular problem; the modeller may need closer control.

2.8.2 Discrete Event Simulation

A discrete event strategy which translates a model into a program is applicable to continuous or discrete time models in which state changes occur at variable but predictable time intervals. Fundamental to discrete event techniques is the formation and management of a "next event list" which orders tasks with respect to time and allows discretization of the time base into the appropriate time steps. Such models assume that nothing happens to component state variables except at these time steps, so that discretization error between model and computer respresentation is not a factor. The problem in maintaining morphisms between model and program in the discrete event strategy is that of ordering and executing the "next event list" in a manner which preserves the interdependence of model components.

Three basic strategies for discrete event programming are in common use (Fishman, 1973). For each strategy packages are commercially available.

The "event scheduling" strategy involves recognizing an

University of Alberta

event as a change in the state of any model component. The "next event list" is initialized with an event which is carried out (the state change implemented), and any consequent changes in state of any component entered in the appropriate place on the list. Conflicts arising from events scheduled at the same time must be resolved with regard to dependence of model components for the particular model. The simulation languages which implement this technique are GASP and SIMSCRIPT which are both FORTRAN based. They may be used when a model suggests the event scheduling strategy.

The "activity scanning" strategy involves recognizing as entities all activities in the model which can affect the states of components, and associating with each a set of conditions (on model component states) under which initiation or termination of the activity occur. For each activity there is a clock, or sequence of time steps at which state changes occur. Processing proceeds by checking the clocks of all activities and advancing time to the most imminent. The activity scanning technique has not been widely used, possibly owing to the fact that a simulation language using this approach is not widely available.

The "process interaction" strategy involves recognition of a process as an entity, where a process is a sequence of events changing the state of a single component. Execution of a particular process may have to be halted to allow contingent execution of other processes in order to preserve the interdependence inherent in the model. In this category

University of Alberta

there are two widely used simulation packages available, GPSS (written in IBM/360 assembly language) and SIMULA (written in ALGOL).

SIMULA is an ALGOL-based simulation language which allows the conceptual operation in parallel of blocks of ALGOL programs. As well as controlling execution of these blocks by using the process interaction technique, it allows description and generation of processes during simulation. The modeller describes an activity procedure for each component of the model, and SIMULA controls interaction and contingent processes in the resulting simulation. SIMULA is a powerful tool where applicable.

University of Alberta

2.9 Summary

In this chapter an outline of the theory and conceptual elements in the process of constructing models and simulating them has been presented. Models and their computer implementations can be presented in a formal language. Within this framework, the success with which one system is represented by another can be formally assessed. Models can moreover be viewed in the perspective of classes of models and facilities available for conversion to iterative computer form.

The remainder of the thesis will be presented in a format corresponding to the five elements, the real system, the experimental frame, the base model (or rather models), lumped models, and computer implementations. Questions of

modelling validity will then be considered in the context of simulation experiments.

University of Alberta

CHAPTER THREE

The Real System

University of Alberta

Since the discovery of the genetic code and understanding of the mechanisms of its translation into chemical activity in the cell, a major area of study for geneticists and developmental biologists has been that of differentiation and development. Scientists have been seeking to discover how cells of common ancestry, with the same information stored in their DNA, proceed to interpret that information differentially, and thus develop into distinct cell types in the cooperatively structured organism. The problem as presently conceived is that of discovering how a cell has in active interpretation at a given time only those portions of its DNA relevant to its present function, and how its present situation leads to other portions of the DNA being active in later progeny cells.

Theories of development have both guided experimentation and been illuminated by it. It is inevitable therefore to discuss such theories (models) at the same time as describing the biological systems themselves. Thus the background of biology and existing theory, out of which the author's own models arise, is presented as the "real system".

3.1 Genetic Information and Development

The programmed development of the multicellular organism seems to be the result of an extremely complex, multilevel control structure. Various experimental techniques attack particular levels in the hierarchy of control.

At the single-cell level, bacteria have been extensively studied, and chemical mechanisms of control over their relatively minute store of DNA have been identified. Portions of the DNA code for proteins, while other portions are sites involved in controlling production. In the well-known 'operon' model, in order for certain code to be read for active production, a portion of code before it, known as the operator, must be in a suitable state. By binding to the operator, certain molecules obstruct, or in other cases promote, reading of following code. In addition, reaction with some other substance may prevent binding by these molecules. Molecules which affect control areas of the DNA may either be extrinsic to the bacterium (thus its

University of Alberta

metabolism can adjust to its chemical environment), or intrinsic, that is, they themselves are coded for in the DNA. In this way the production of one metabolite can affect production of another, or a critical concentration of a metabolite can affect its own production. In the bacterium, therefore, a picture emerges of control at the chemical level, with several types of feedback.

In multicellular organisms DNA organization is thought to be immensely more complicated. The quantity of DNA is orders of magnitude larger. Cellular activities are correspondingly more complex, and the identification of specific chemical mechanisms in control becomes very difficult. Geneticists using mutation experiments, however, have demonstrated that only a fraction of the DNA can be directly implicated in the production of detectable substances in an organism. Furthermore, some mutations cannot be associated directly with these structural regions of the DNA. Since such mutations frequently have serious consequences, they are perhaps the result of changes in control portions of the DNA which affect future development. It has also been demonstrated that those portions of the DNA which appear to be implicated in control contain reiterated code. It has been suggested that such repetition could be a means of timing events in the balanced, self-regulating development of the organism.

A model showing possible levels of control coded for in the DNA of a multicellular organism has been proposed by

University of Alberta

Britten and Davidson (1969). In this model, batteries of "producer" genes are activated by cascading levels of controlling genes which in turn are initiated by the product of a single "sensor" gene. Different batteries can interact at levels below the sensor gene, if products of genes in one battery affect the activity of genes in another.

The analysis of DNA structure does not give a complete appreciation of the mechanisms of development and cell differentiation, however. Development is above all a dynamic process. The DNA forms the data for a program which accesses that data in a sequence dependent upon the results of the part of the program already carried out. The DNA of a cell with its "active" and "inactive" portions may be thought of as the cell's "state". Development is the action of a "transition function" taking a cell and its progeny from one state to another. In order to describe the process, one of two methods can be used. In the first, a "bottom-up" method, the following steps would be taken:

(a) all elements in the state of a cell would be identified; that is, all of the relevant functional products of the organism would have to be listed;

(b) all external influences on (inputs to) a cell would have to be identified;

(c) a "snapshot" of the organism for each state occurring in all individual cells would have to be taken;

(d) with all of the above data, the nature of the transition function, and thus of the developmental process, could

University of Alberta

(possibly) be inferred.

The bottom-up method described above is the classic means of obtaining an operational, dynamic picture of a process from successive static descriptions. It is obviously an extremely limited approach, both practically and conceptually, in the case of complex multicellular organisms. For that reason experimental techniques and control models of a "top-down" nature have developed. In this second approach, the following steps would be taken:

(a) a developmental phenomenon would be identified in the organism being studied;

(b) any regularly occurring events, or physical patterns, which might be associated with the phenomenon would be observed;

(c) experiments designed to disturb such events or patterns would be performed in order to establish possible causal relationships;

(d) once such causal relationships had been established, results would be compared to others obtained in developmental studies to test their generality and plausibility.

Two instances of the fruitful use of a top-down approach will be given.

In a well-known embryological phenomenon called induction, when a certain kind of tissue (body of cells with specialized function) grows to within a set distance of another tissue of different type, a change is observed in

University of Alberta

the second cell type. If growth to the critical distance is prevented, no change occurs. The implication is that some product of cells in the first tissue triggers a chain of events in the second type which results in a change in the genetic activation of the second type. In terms of an overall picture of development, such phenomena show that it is important to think in terms of populations of cells and interactions between them.

As a second example, it has been shown that the position of a cell in a Drosophila imaginal disc (see Section 3.2) determines the cell's future. The implication is that cells have some means of measuring their position in a population. Again the significance of populations of cells and signals between them is indicated.

The strengths of a top-down approach are that it does not attempt to look at all of the cell's possible activities at once, but concentrates on a certain sequence of events, that it can keep as many cells (as much of the organism) as desired in view, and that since it concentrates on change it is dynamic in approach.

It should be emphasized that the top-down type of approach advocated here is a "difference" method. It does not try to describe a cell's state of activation in entirety but rather characterizes differences between cells. Studies of induction are not concerned with all products of the cells but only those which are changed by the inductive event. The characterization of populations of cells in

University of Alberta

terms of their developmental history, then, can be given in terms of changes in behavior. These changes and differences between cells accumulate as development progresses. If all differences between all cells at different stages in development were to be catalogued, together with the influences causing them, a complete description of the genetic apparatus of the cell would be achieved. The power of a difference method is that it proceeds from simplicity (cells which are initially not mutually differentiated) to complexity, in a dynamic manner, without trying to maintain a universal perspective.

University of Alberta

3.2 Disc Studies in Drosophila

In the larval stage of development of the fruit fly Drosophila small buds of tissue called imaginal discs form. They are destined to become the external limbs and organs of the adult fly. These discs can be manipulated experimentally, cultured in the abdomens of adult flies, implanted back into larvae and allowed to differentiate during metamorphosis. The results of the experimental manipulation can be observed in the abnormal structures resulting in the mature host fly. From the extensive research on imaginal discs, a few pertinent results will be described here.

3.2.1. Transdetermination and Codes

Hadorn (1966) has shown that discs undergo

determination with respect to their future organ type at an early stage, long before visible expression of this type. All progeny cells inherit the same determination as their parent cells. Hadorn found that with repeated culture of discs in adult abdomens, a small fraction of these discs change their state of determination and become forerunner cells for another disc type. This phenomenon is known as transdetermination. The frequency of transdetermination between the various disc types was calculated for aabout 6,000 implanted fragments observed through metamorphosis.

Kauffman (1973) has used Hadorn's transdetermination data as the basis for a model of determination in terms of a number of bistable switches. The idea is that transdetermination occurs because the disc-type decision is of a switch-like nature, and decisions can be changed by resetting switches. Kauffman's assignment of bistable switches is shown in Figure 3.1.

The frequency of transdetermination should be inversely proportional to the number of switches which are in a relatively stable state. A minimum of 4 switches is required for good agreement with the frequency data. The 1 state of a switch is assumed more stable than the 0; thus frequencies of transdetermination in opposite directions are not equal. For example, Eye to Wing happens more often than Wing to Eye, as suggested by the lengths of the corresponding arrows.

Kauffman's model is a difference model of the type

University of Alberta

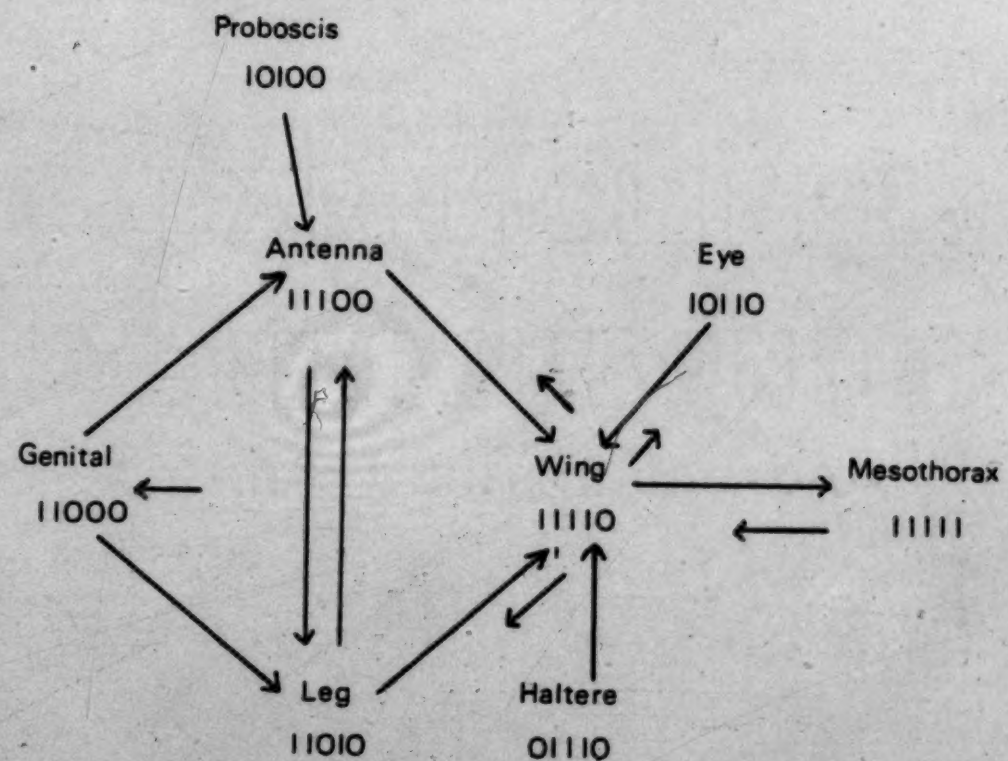


FIGURE 3.1 Kauffman's assignment of bistable switches

discussed in the previous section. It is not a picture of activation and de-activation in the DNA corresponding to disc type, but rather a minimal description of differences among cell types and transformations between them. The model is important because it expresses these differences in quantitative terms. Hamming distance between the codes for different discs has meaning in terms of genetic similarity between them.

3.2.2. Positional Information and Gradients

A number of experiments indicate that the location of a cell in a disc has more significance than determining the final structure of the organ. For example Schubiger (1971) has found differences between parts of the disc in terms of their capacity to produce duplicates of the features they determine, when fractured from the rest of the disc and allowed to grow, or to reproduce features which were removed. He has also found a range of frequencies of transdetermination for different fragments. So, before any question of final differentiated cell type arises, cells have some way of sensing where they are in the disc; that is, the cells possess positional information.

Bryant (1971) has developed the idea of an information gradient in cells. In situ cuts of discs into two pieces result in the regeneration of the entire disc from one portion and mirror image formation by the other, judged by the regenerative or duplicative effects seen in the adult

University of Alberta

organism. This phenomenon can be explained in terms of a gradient of regenerative potential, as shown in Figure 3.2. The portion of disc containing the regenerative hotspot A will regenerate, while the other portion can only produce features lower than its own highest point on the gradient. Region A has been located on the disc.

A large amount of evidence points to this capability of cells to read information of a positional nature. As one example, if a portion of the developing leg of a cockroach is removed, compensatory growth occurs. This can be explained in terms of the existance of a gradient down the leg; when a portion is removed, a discontinuity is caused in the gradient which acts as a signal to cells to divide in that region.

French, Bryant and Bryant (1976) have developed a formal model for pattern regulation, using polar coordinates to describe a gradient field in two dimensions. There is direct evidence for two simple rules governing regeneration in cockroach legs, imaginal discs and amphibian limbs. The "shortest intercalation rule" stipulates that when normally nonadjacent positions, either in the radial or the circular sense, are confronted in a graft combination or in wound healing, growth occurs at the junction until all intermediate positions have been intercalated. (If a circle of positions is interrupted, missing positions on the shortest arc joining the confronting cells are intercalated.) The "complete circle rule" stipulates that an

University of Alberta

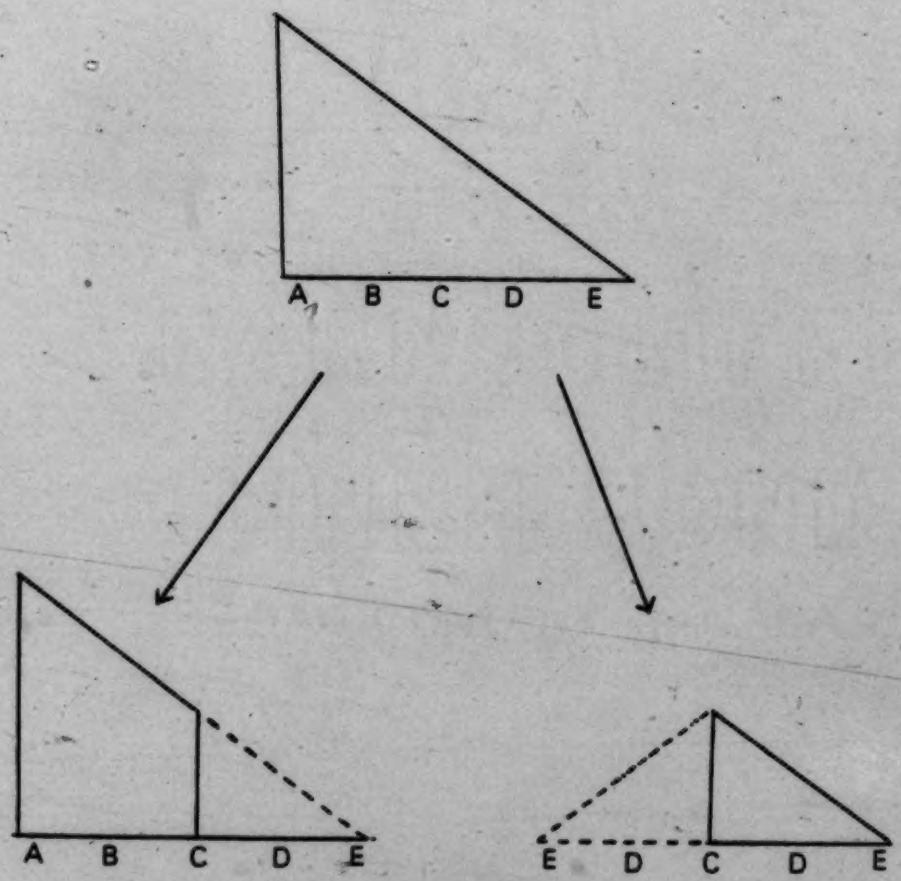


FIGURE 3.2 Bryant's gradient model

entire circle of positional values, at a certain radial position, may undergo distal transformation to produce cells with all the more central (distal) positional values. That is, the circle will produce all values contained in a disc extending to that radial position. The authors suggest that since the model can explain phenomena in diverse organisms, it may be generally applicable to pattern formation.

Several questions are open concerning gradients. Some of these are:

(a) What is the physical basis for these gradients? Are they as simple as concentration gradients?

(b) How many different fields of positional information are needed to give a cell its future developmental instructions?

(c) What shapes of fields exist? Are they rectangular, spherical, or perhaps spiral chemical waves?

University of Alberta

3.2.3. Compartmentalization

Work of great interest has been done in the last few years in the laboratory of Garcia-Bellido and his co-workers (1973). An excellent summary of this work can be found in a paper by Crick and Lawrence (1975), who indicate the importance of the results to developmental questions.

A mutation in a particular cell is called 'gratuitous' if it does not interfere with the normal development of the organism. If a gratuitous mutation results in progeny cells which are identifiable against a background of non-mutant cells, on the epidermis for example, a technique known as

clonal analysis can be used. In this technique lineages of cells are followed by studying the descendants of a number of marked cells in different specimens. Garcia-Bellido et al. marked cells in this way by using irradiation of the wing disc in a special genetic strain of Drosophila.

The earlier a cell is marked in the developing disc, the larger the number of marked descendants there will be in the observable mature wing. The ratio of marked to unmarked cells indicates how many cells there were at the time of irradiation. If there were n cells at irradiation, and one is marked, then 1/n of all cells in the mature wing will be marked. The patches of marked cells were initiated at various times in the development of the disc. They show coherence, and the startling phenomenon which has come to be known as compartmentalization. Marking cells one generation apart leads to the discovery that the disc undergoes a number of subdivisions into distinct areas with boundaries impervious to intrusion by progeny cells from bordering areas. A cell marked just before such a division may have progeny in both compartments while a cell marked after the division does not. If the clone of marked cells on the wing encounters a boundary the patch follows it. Thus it has been discovered that the lines of division are smooth and invariant from specimen to specimen. Furthermore, the division events occur in a set sequence and are inescapably implicated in the developmental program of the organism.

Investigations in other discs, and other organisms,

University of Alberta

have since shown that compartmentalization is a frequent event in development. It may be even more frequent than can be seen because it is readily recognisable only in systems of a convenient two-dimensional nature such as the epidermis of a wing.

Crick and Lawrence see three possible mechanisms for the establishment of compartmental boundaries.

1. Partition of daughters: each cell has one daughter which ends up in one compartment and one in another.
2. Random allocation: cells are allocated to compartments at random according to some statistical measure keyed to the size of the resulting compartments.
3. Geographical partition: founder cells of one compartment are separated from those of another by the establishment of a physical line of division.

The first two mechanisms seem unlikely, as a large amount of cell movement and sorting out would be necessary to allow grouping of the cells of a compartment. That much movement is unlikely in view of the coherent clones of marked cells which have been observed. The third mechanism appears the most likely, and would be the result of a phenomenon of field-like nature. Gradients of information across the disc again appear necessary.

Crick and Lawrence propose a list of questions stimulated by compartmentalization studies. They are summarized below.

(a) Are all divisions binary? If they are, does this imply

University of Alberta

a binary switching mechanism in the DNA itself?

(b) Why are compartment boundaries smooth, even to within a cell diameter? Do boundaries lie along smooth contours of fields of information, or is there cell sorting between cells of distinct characters, or a combination of the two mechanisms?

(c) When do subdivisions cease? Does this kind of phenomenon continue until final cell differentiation itself, or are there other mechanisms at work?

(d) How are cells in different compartments actually different? Are there immediately expressed differences as well as final differentiated cell types and organ parts?

(e) What is the connection (if any) or conflict (if any) between compartments and gradients of information?

(f) Are compartmental decisions irreversible? How stable are genetic decisions?

(g) What is the significance of the coincidence of compartmental boundaries with regions of uniformly transformation in homeotic mutants (mutants in which one tissue is uniformly transformed into another in a specific region) (Kauffman, 1978)? Can the homeotic genes be identified with Britten and Davidson's (1969) sensor genes?

In a subsequent paper Lawrence (1975) has attempted to discover the degree of discontinuity across compartment boundaries. The cells appear to be functionally distinct in

University of Alberta

some but not all respects. Communication of a chemical nature across boundaries is possible but may be selective.

In summary, compartmentalization appears to be an essential part of development in many organisms. It appears to be the basis for organization of populations of cells into functionally distinct groups and populations of differentially communicating cells. The connection between compartments and gradient fields of information is an important focus of attention in present studies. One of the major difficulties, of course, is that the nature of gradients remains largely obscure.

Three fundamental ideas have been presented in this discussion of issues in genetics. They are characterization of the problem in terms of difference models of a quantitative nature, the importance of positional information in development, and the phenomenon of compartmentalization.

University of Alberta

3.3 The Mechanics of Pattern Formation

A fundamental process in development is the formation of spatial patterns of differentiated tissues from initially homogeneous tissue. Individual cells in a specific tissue have differentiated, but formation of the spatial structure of the tissue necessitates a super-cellular mechanism. It has long been thought (Child, 1941; Waddington, 1962) that "primary fields" of morphogen concentrations, or of other spatially varying physical parameters, must be established

as pre-patterns for tissue structures. The genetic constitution and history of a cell provides specificity of response to positional signals; differential response of cells of similar genetic constitution depends on positional specificity extrinsic to the genome (Garcia-Bellido, 1975). Thus topological organisation precedes and conditions specific gene activation.

Investigations of various intercellular mechanisms capable of establishing spatial patterns have been undertaken. These investigations are typical of the "top-down" approach of the developmental geneticist in that specific interaction of morphogenetic agents and the DNA is in many cases unknown. However if spatio-temporal patterns of such agents are found to coincide with developmental structures, a causal (or at least influential) relationship may be inferred.

A morphogenetic field may be defined as an assembly of functionally coupled cells whose development is under the control of a common regulatory process. Establishment of a pre-pattern in such a field implies a mechanism by which an initially homogeneous distribution of a morphogenetic agent evolves a spatial inhomogeneity capable of initiating and sustaining a pattern of differentiating cells. Several such mechanisms have been studied.

University of Alberta

3.3.1 Active Transport

If a polarized asymmetry exists in each of the cells of the field, a morphogen can be actively propogated in a particular direction through the field, resulting in a gradient of concentration. Boundary cells in such systems have special properties sustaining the gradient (Lawrence, 1966; Wilby and Webster, 1970). A similar system involving boundary "sources" and "sinks" and unpolarized cells, has been developed by Babloyantz and Hiernaux (1974). Inhomogeneity in such models is a result of a mechanism which actively opposes the natural diffusion of molecules down their concentration gradient to achieve uniform concentration, the minimum energy configuration.

3.3.2 Phase Gradients

If the cells of a field can be regarded as autonomous oscillators which emit signals of short duration at regular intervals, and the signals of a cell are propogated by other cells after a delay, a wave of concentration from each cell arises. If a gradient of emission intervals is pre-existent in the field (implying an initial polarity), a stable pattern of positional information can result (Goodwin and Cohen, 1969). Aggregation in cellular slime molds is an example of the organisational capability of oscillatory processes (Nicolis and Prigogine, 1977).

University of Alberta

3.3.3 Cellular Contacts

Positional specificity of morphogen concentration can be obtained from an initially polarized field without widespread propogation if the concentration level in a cell is affected by the neighbour in the polar direction (Babloyantz, 1977) or neighbours in bipolar directions (Garay, 1977).

3.3.4 Cell Sorting

If it is assumed that the genetically identical cells of a field are in some sense differentiated with respect to some adhesive or chemotaxic property, then conceivably a cell-sorting mechanism could give rise to the assembly of patterns of cells by autonomous movements of cells in the field.

All of the above mechanisms involve either initial polarity of the field or initial asymmetry of cells, or both, and thus violate strict interpretation of the initial homogeneity of the field. It has been shown that processes naturally occurring in fields of cells, which initially achieve steady state (time invariant), spatially uniform conditions, may, under certain influences also naturally present, deviate from such conditions and produce a non-uniform stable state. An initial investigation of such situations was undertaken by Turing (1952). Recently several theories of pattern formation based on similar phenomena have been formulated (Gmitro and Scriven, 1966;

University of Alberta

Crick, 1970; Gierer and Meinhart, 1972; Meinhart, 1977; Kauffman et al., 1978).

3.3.5 Reaction-Diffusion Kinetics

Consider a field in which local conditions are defined by the concentrations of various chemical species. Any irregularities in these concentrations tend to be damped by diffusion; that is, concentrations tend towards spatial uniformity by transport of the species down its concentration gradient. Diffusion may be regarded as a consequence of the second law of thermodynamics, which states that in a closed system entropy increases monotonically until its maximum is reached at "thermodynamic equilibrium". Thermodynamic equilibrium in the field corresponds to steady state, spatially uniform concentration.

Irregularities in concentrations are caused either by chemical reaction within the field or by mass exchange from the surroundings. For the concentration of a chemical species X(r,t), the following reaction-diffusion equation can be established:

$$d(x)/d(t) = F(X,Y,Z,...)+D(1)L(X)+M(x) \qquad (1)$$

That is, the concentration evolves in time according to: (a) chemical reaction, F(X,Y,...) (which may involve other species Y,Z,...); (b) diffusion, which may be approximated in dilute mixtures by an independent, constant diffusion rate D(1) multiplying the Laplacian operator (sum of second

University of Alberta

spatial derivatives, or rate of flux of X at r); (c) mass influx through the boundaries M(X).

An equation of this type is classified as a parabolic problem in that the time derivative is of first order and spatial derivatives are second order. The spatial pattern of concentration of X will depend on the stability of the thermodynamic equilibrium, steady state, spatially uniform solution given by F(X0,Y0,Z0,...)+M(X0)=0. The term M(X) may be considered as introducing boundary conditions; M(x)=0 on the boundary corresponds to a closed system with no flux across the boundary (a "Neumann" condition), while M(X) of a certain value on the boundary (a "Dirichlet" condition) corresponds to an open system (for example, "source" and "sink" problems). Specified boundary conditions for the reaction-diffusion equation ensure the uniqueness of the steady state solution X0.

It has been shown (Turing, 1952; Nicolis and Prigogine, 1977) that for certain types of reaction functions F(X,Y,Z,...) the thermodynamic equilibrium solution X0 is not stable, and perturbations produce stable, spatially inhomogeneous patterns of concentration. Some analytical techniques described below may be used to examine this situation.

1. Bifurcation Theory. Bifurcation theory was initiated by Poincare and developed by other mathematicians interested in non-equilibrium solutions of nonlinear differential

University of Alberta

equations (Nicolis and Prigogine, 1977).

Suppose in equation (1) that M(X)=0 and F is governed by a set of parameters p (corresponding to rates of reaction). If the p are small enough, solution of the system evolves towards thermodynamic equilibrium (X0) as t tends to infinity. Now for critical values of p the system can evolve to a new steady-state solution due to instability of the thermodynamic equilibrium. Such values of p are called "bifurcation points". The purpose of the theory is to develop methods to (a) locate such points for a given system and (b) construct approximate analytic, convergent expressions for solutions emerging at bifurcation points. An example of the application of bifurcation theory to a reaction-diffusion system may be found in Nicolis and Prigogine (1977).

University of Alberta

2. Theory of Catastrophes. In further development of the work of Poincare, Thom (1972) has developed a means of classifying systems which can undergo transitions from one steady state to another. Such systems must be expressible as first order ordinary differential equations with time as the independent variable. The diffusion term in (1) is not present explicitly; all spatial dependence is expressed by a parametric dependence of F in space and time. The time derivative arises from a "potential" function

$$dX/dt = - dV(X,Y,Z,...)/dX \qquad (2)$$

Steady states of the system correspond to dV/dX=0 and exist

if V does not depend explicitly on t. dV/dX=0 defines hypersurfaces in the parametric space. Identification of points where there may be transition between steady states becomes a topological problem of identifying regions in which a curve (particular solution) may cross from one hypersurface to another. An example of the application of catastrophe theory to a one species reaction system may be found in Nicolis and Prigogine (1977).

The potential importance of catastrophe theory in describing phenonema of biological order has been emphasized by Goodwin (1973). Severe criticisms of some attempted applications have been made, however (Zahler and Sussman, 1977); it appears that a fundamental problem arises out of the present lack of quantitative guidelines to its practical application in problems of development.

3. Stability Theory. Behavior of a reaction-diffusion system such as (1) under perturbation of the equilibrium solution may be studied by investigation of the stability of that equilibrium configuration X0. Stability can be investigated by "Lyapunov's Second Method" (Nicolis and Prigogine, 1977), in which a functional must be constructed of the concentrations. The functional plays the role of a potential in the vicinity of X0. Asymptotic stability at X0 is established if a (local) absolute minimum of V exists there. Asymptotic stability is the property that any solution in a small neighbourhood a of X0 at time t(0)

University of Alberta

approaches X0 as time tends to infinity.

The stability of X0 may also be investigated by means of linearization of the system in the neighbourhood of X0. This method has been investigated in detail by Gmitro and Scriven (1966), is discussed by Nicolis and Prigogine (1977), and has been practically applied to specific problems in pattern formation by Turing (1952), Meinhardt (1978), Wilby and Ede (1976), Kauffman et al. (1978), and others.

The steady state (time-invariant), spatially homogeneous (diffusion term zero) solution X0 of (1) obeys

$$0=F(X0,Y,Z,...)+M(X0) \qquad (3)$$

Unsteady states can be represented by the sum of steady state concentration X0 and a depature from the steady state, x; that is, X=X0+x.

Equation (1) can the be written

$$d/dt\ (X0+x)=F(X0+x,Y,Z,...)+D(1)L(X0+x)+M(X0+x) \qquad (4)$$

Under approppriate smoothness conditions, F and M can be expanded in terms of a Taylor's series in x around X0; further, if x is small enough, terms of order x**2 and higher may be neglected in these expansions. Then (4) becomes

$$dX0/dt+dx/dt=F(X0,y,z,...)+dF(X0)x/dx+D(1)1(X0)+D(1)1(x)+ M(X0)+M(x) \qquad (5)$$

By (3) and dX0/dt=0 this becomes

$$dx/dt=dF(X0)x/dx+dM(X0)x+D(1)L(x)$$

University of Alberta

or $dx/dt=Kx+D(1)L(x)$ (6)

This is a linear equation which may be solved by the methods of harmonic analysis. For several species X(1),X(2),X(3),... whose reaction-diffusion equations are coupled, K is the Jacobian matrix of F+M with respect to x(1),x(2),x(3),... evaluated at the thermodynamic equilibrium solution of the system of equations.

The methods of harmonic analysis are based on the fundamental theorem of Fourier which states that any spatial configuration in a domain may be expressed as a weighted sum of elementary shape functions appropriate to that domain. The weighting factors relate the time-dependent pattern particular to x to the elementary patterns. Thus (x) (the column vector(x(1),x(2),...,x(n))) may be written

$$(x)=(a(k,n)(A)(k,n)\exp(L(k,n)t))F(k,r) \quad (7)$$

where

a) F(k,r) are the shape functions of the domain

b) L(k,n) and (A)(k,n) are the eigenvalues and eigenvectors of the matrix K-k**2D, D=D(n)I

c) a(k,n) are constants determined by initial conditions.

The eigenvalues and eigenvectors are found by solving

$$\det(K-k^{**}2D-1I)=0 \quad (8)$$

The ability of the system to exhibit pattern formation depends on these eigenvalues, which depend in turn on the rate coefficients in K, the diffusion coefficients in D, and

University of Alberta

the "pattern size factor" (or spatial eigenvalue) k. The eigenvalues are in general complex; the real part is a growth factor and the imaginary and oscillatory frequency associated with the pattern. If any eigenvalues have positive real parts, the system is said to be unstable with respect to pattern of size s=2/k, because any disturbance of the steady state containing an element of this will grow exponentially. If all eigenvalues have negative real part, the system is stable.

While the departure from equilibrium remains small enough to ensure correctness of the linearized stability study, the pattern size for which the growth factor is largest will dominate the disturbance. Thereafter nonlinear effects come into play (Gmitro and Scriven, 1966). If the eigenvalue having largest real part is entirely real, a "standing wave" of concentration, in which concentration at each point steadily increases, will be seen; if it is complex, a "travelling wave" will be seen, in which concentrations at each point cycle with frequency given by the imaginary part and grow in amplitude.

In summary, a stability analysis of a reaction-diffusion system shows the following results. A spatially uniform steady state of the system is subject to small disturbances, whether from external influence or internal "thermal noise" or "molecular fluctuations". The stability of the steady state in response to such disturbance depends on the reaction and diffusion coefficients of the system; if

University of Alberta

these are such that instability ensues, spatial patterns will manifest themselves. The nature of these patterns depends on the spatial domain of the system and the eigenvalues of K-k**2D.

Gmitro and Scriven (1966) give a detailed analysis of the patterns which may arise in a line, a circle, a cylinder, a ring, a plane (both in rectangular and polar coordinates) and a sphere. That is, they give the eigenvalues or elementary patterns for the Laplacian for each domain, and the values of k**2 when, in a closed domain such as a circle with conditions imposed on the boundary, it must assume discrete values in order for patterns to emerge. They undertake a linearized stability analysis for one, two or three participating species (corresponding to the number of coupled reaction-diffusion equations).

University of Alberta

3.3.6 Applications of Reaction-Diffusion Kinetics

A brief description of instances in which reaction-diffusion kinetics have been used in modelling pattern formation is given below. Some of these models will be discussed in more detail when computer-oriented models are described.

Turing (1952) investigates the departure from equilibrium of linear reaction-diffusion systems by means of linearized stability analysis. Linear systems operating in a ring of cells or sphere are demonstrated to develop standing or travelling waves of concentration of

exponentially growing amplitude; a discussion of gastrulation in a mammalian embryo is linked to wave formation in a sphere. However in linear systems no limiting of the exponential growth factor is possible and thus no new stable state or pattern can be established. Nonlinear terms coming into play as the system gets further from the equilibrium state appear to balance this exponential growth of amplitude in the wave pattern and allow a stable pattern to be established (Gmitro and Scriven, 1966; Nicolis and Prigogine, 1977; Grierer and Meinhart, 1972).

Crick (1970) has investigated the feasibility of reaction-diffusion as a means of pattern formation in embryogenesis. Assuming a typical size of a primary field as O(10) (between about 30 and 100) cells, and the establishment of the field (pattern) within 5-10 hours, he estimates the size of a cell and rate of diffusion necessary if diffusion is to be assumed the means of spatial communication of intercellular differences in concentration. He obtains a value for the diffusion rate which is of the same order as that of a typical metabolite (c-AMP), and a cell size of 10-30 micrometres, which is realistic. The conclusion is therefore that the parameters of the diffusion process are such that it is a practical mechanism for establishment of primary fields.

Grierer and Meinhart (1972) and Meinhart (1977) use two coupled, nonlinear reaction-diffusion equations for

University of Alberta

activator and inhibitor species in modelling pattern formation in organisms which may be regarded as one-dimensional. The models assume initial gradients of the density of cells which are sources of the two chemical species. Then, if a reaction-diffusion system is set up in which the activator (which catalyzes both itself and the inhibitor) has a lower rate of diffusion than the inhibitor (which inhibits only the activator), patterns of concentration of the two species are established with an accentuated peak at one end of the one-dimensional array. It is suggested that such patterns could act as prepatterns for long-term differentiation of additional cells into sources of the species. Computer implementation of the model (by unspecified numerical techniques) is compared to quantitative data from hydra, and suggests that the model can account for certain duplication phenomena.

Use of an initial gradient of density of source cells (which are cells differentiated from the general population) in order to obtain a pre-pattern for further cell differentiation suggests that the model is one of pattern "reinforcement" rather than pattern "formation". Establishment of the initial density gradient has not been explained. Less initial disparity is necessary in the work of Meinhart (1977), where the field is polarized by a single source of activator and inhibitor at one end. Such a model is used to discuss segment formation in insect embryos.

A sophisticated application of linearized stability

University of Alberta

analysis of a two-species, nonlinear reaction-diffusion system in explaining the sequential formation and shapes of compartments has been proposed by Kauffman et al. (1978). In their model, the temporal eigenvalues of the linearized R-D system are related to the size of the growing disc, which they propose to be approximately elliptical. The steady-state, spatially uniform solution of the system pertains except when the disc attains a sequence of critical sizes. At these sizes, a particular eigenfunction of the Laplacian (which in an ellipse are the Mathieu functions) fits exactly into the ellipse obeying a zero-flux boundary condition; thus it is postulated that at such critical sizes the steady-state, spatially uniform solution becomes unstable and a pattern corresponding to the favoured eigenfunction is established. Such a pattern could act as a prepattern for a compartmentalization event, with a compartment boundary specified by a threshold value of one of the species.

As the disc grows larger, the pattern no longer "fits" and decays to the uniform situation, until the next critical size is reached and the next pattern established. The Mathieu functions (axes of the ellipse, confocal hyperbolae, and confocal ellipses) can be made to agree in general shape and timing with compartmentalization in the wing disc.

Reaction-diffusion relationships of a primitive kind have been used in signalling between cells in computer-oriented models, such as the one dimensional model of Hydra

University of Alberta

developed using CELIA (Herman and Rozenberg, 1975) and other models based on L-systems (Lindenmayer, 1976). Such modelling of intercellular signalling has been in the form of uncoupled, linear difference equations and will be discussed at greater length in the next section. In general, the methods fail to investigate the numerical stability of the difference equations employed. In consequence, even if the equations were of sufficient complexity to admit pattern establishment of the kind under discussion, the numerical results obtained could not demonstrate them.

Wilby and Ede (1974) have developed a model and simulation for differentiation of muscle tissue into bone in the growing chick wing. A cell-to-cell signalling system in terms of difference equations involving one species undergoing synthesis, "diffusion", and destruction that is dependent on two thresholds (for initiation and termination) is set up to obtain a periodic pattern of concentration down the limb. Such a pattern could act as a pre-pattern for the skeletal pattern of the limb. The authors suggest that such a periodic distribution is more plausible than a situation in which a monotonic gradient model necessitates interpretation of a series of thresholds of different magnitudes being interpretated to result in simular cellular events (formation of bone cells). The difference equations are designed to produce such periodicity; it may be pointed out that the eigenfunctions of the Laplacian on a rectangle

University of Alberta

(corresponding roughly to the shape of the limb at intermediate stages of development) are trigonometric in form, and thus use of a reaction-diffusion system should allow the natural establishment of a periodic pre-pattern.

If complicated structures such as the limb skeleton are to result from simple, non-periodic morphogen gradients, then one mechanism might be the interaction of several such gradients. An example of such a model is given by MacWilliams and Papageorgiou (1978), who introduced a pre-pattern for the chick wing specified by thirteen morphogens each of exponential distribution. Such a model can be constructed from the fully developed limb by regarding it in the light of a two- or three-dimensional map of contours around the various structures. The shape of contours around a feature determines how many gradients would be needed to generate such a field. The authors suggest that cellular differentiation at such a feature is determined by a "winner take all" rule on binding the morphogen of largest concentration to the genome. The model is complex and teleological in construction; furthermore, the establishment of the numerous gradients is not discussed. In contrast, a reaction-diffusion system with few interacting species is capable of establishing equally complex patterns in a comprehensible manner.

Grossberg (1978a,b) draws an analogy between the stability characteristics of reaction-diffusion systems (particularly those involving coupled, long term inhibition

University of Alberta

and short term excitation such as in Grierer and Meinhart, (1972), and Kauffman et al. (1978)) and the stability characteristics of shunting networks where local interaction in governed by ordinary differential equations. In both types of systems dynamically changing patterns can emerge from initially uniform conditions. Successful approximation of reaction-diffusion systems in terms of finite differences on a grid produces a representation mathematically equivalent to serial, discrete operation of mass-reaction laws in a network of the type presented by Grossberg. Grossberg's approach has been directed towards models in population biology, ecology and psychophysiology, with some discussion of parallel phenomena in developmental biology. He suggests a fundamental class of control mechanisms in seemingly competitive systems in which pattern or order appears apparently spontaneously.

Various investigations have been made, therefore, of ways in which biological systems, and in particular primary fields, can develop pattern by the operation of small scale, chemically feasible mechanisms. The success with which such mechanisms can be modelled with the help of the computer will be discussed in succeeding chapters.

3.4 Computer Oriented Models

The "real system" has been presented as a mixture of biology and existing models of biological phenomena. A number of models of biological system in which simulation

University of Alberta

plays an essential part have been constructed, and are described below.

3.4.1 Models Based on L-Systems

Developmental algorithms for multicellular organisms have been developed based on the work of Lindenmayer on L-systems (Lindenmayer, 1975). An L-system is a finite automaton in which components of the array can be replaced by arrays. Because of the complexity of the geometry of such replacements in two and three dimensions, most of the successes of L-systems have been in the modelling of structures of a one dimensional character. Extension to higher dimensions by means of "graph-L-systems", where components are nodes which can be replaced by nodes and links, is being attempted (Lindenmayer, 1975).

An L-system is a discrete/discrete model: components (cells) are discrete, and the state set is discrete. The state of a cell usually refers to a genetic state which is heritable and may change according to a transition function which is uniform over the array, the present state of the cell, and the states of its neighbours. The transition function may be deterministic or non-deterministic.

One-dimensional L-systems correspond exactly to formal language theory: the set of states is an alphabet, and all developmental possibilities may be generated by the operation of the transition function (set of productions) on the starting state (axiom). Because of this formal

University of Alberta

equivalence, many predictions about L-systems can be made theoretically. Systems are classified according to the connectivity, the number of neighbours affecting each cell, and have been used to model appropriate developmental systems of a filamental nature.

The work of Herman et al. (1974) applies the techniques of L-systems to development in simple two-dimensional organisms. Their FORTRAN program CELIA simulates linear iterative networks of identical modules (or cells). Cells have a finite number of possible states and change state depending on their own previous state and threshold values of signals from their two neighbours. A complete set of rules for change of state must be supplied by the user. The end cells in a line can communicate with the environment. Cells divide synchronously. Branches can be grown on linear organisms by means of a bracketted list structure.

CELIA has proved useful in demonstrating the effectiveness of a linear gradient in producing coordinated changes of state in simulations of certain organisms. The developmental rules for red algae, patterns on sea snail shells, and grafting experiments in Hydra are some of the areas to which the program has been applied. It has been demonstrated that a polar regulation of cell states can be caused without imposing polarity in individual cells.

The limitations of CELIA are apparent. Only linear gradients parallel to the linear iterative arrays can be

University of Alberta

investigated. The authors themselves point out the need for probabilistic changes of state and cell division. Their simulation is implemented in FORTRAN, making their data structures difficult to maintain. A departure from strict synchronous processing in such structures is introduced by Hogeweg (1978), in which advantage is taken of the low activity level in the array by using a discrete event technique and SIMULA in programming the model.

3.4.2 Growth Models in Arrays of Cells

Wilby and Ede (1976) in further study of the chick wing, have constructed a simulation of growth. A rectangular array with initial cells on one border is subjected to various gradients of mitotic capability. A cell divides probabilistically if its internal situation places it above a threshold of mitotic ability determined by its position. The shortest path to an unoccupied position is then found, and all cells along that path displaced outwards. By these means the shapes of normal and abnormal wings have been obtained, showing that limb shape is possibly mediated only by controlled cell division.

The Wilby and Ede simulation is an elegantly simple model of growth, but intercellular communication and the data structures needed to describe it are not considered.

A simulation of embryonic development by Raven and Bezem (1976) attempts to characterize differences between cells. They represent cells as spheres of different size,

University of Alberta

with cell flattening along contact surfaces represented by the chord of intersection of two spheres. This method permits a close view of the early stages of development, when there are not many cells. By assuming that the type of a cell depends on concentration gradients of axial and radial shape, they produce realistic cell differentiation patterns in a simulation of Lymnaea.

A discrete/discrete model of cell division which is used to study clonal patterns has been developed by Ransom (Ransom, 1977) and programmed in an ALGOL-based language Edinburgh IMP. Cells do not change state (therefore the model is not truly developmental) but are defined by a heritable clonal marker and a position in an array of fixed maximum size. The array is hexagonally arranged. Cells are processed sequentially in each time step; a cell is allowed to divide if it has not divided already in the current time step, and if its division does not cause separation of any pair of cells having the same clonal marker ("stickiness" rule). When a cell divides, the nearest free space is found and cells along that direction are displaced outward. The displacement is simulated by moving the entries for cells in a hash table representing the array.

Cells near the centre of a population are less likely to divide in this model as they are more likely to violate the "stickiness rule" than cells near the edge. Ransom has used his model to simulate growth in Drosophila imaginal discs, in which boundary cells do divide more frequently

University of Alberta

(Ransom, 1977b). Division in the direction of nearest free space implies division in radial directions in a "circular" disc, and Ransom argues that straight radially oriented clonal boundaries observable at some stages of disc development may be produced by such directed cell divisions.

Ransom's model is not genetically oriented; cells do not change state, nor is their state described by more than one discriminant. Furthermore they do not communicate. Moreover cells divide without respect to local conditions or their own state. The only factor affecting division is the "stickiness rule", in which an event remote from a cell may prevent its division.

Duchting (Duchting, 1977) has constructed a model of growth in a 10x10 matrix of cells using the microprocessor system INTEL 8080. It bears some similarity to Conway's Game of Life in that cells become active, die or divide according to the activity status of their neighbours. Cells inherit a marker which also indicates a division rate. The model has been used to simulate cell populations of different division rates in competition, and may therefore be of utility in studying the proliferation of cancer cells.

3.4.3 Intracellular Growth Models

Various models have been constructed which attempt to describe the internal factors of a cell's biochemistry which determine its division cycle time. Simulation of such models results in the collection of statistics about the

University of Alberta

cells in a population such as the distribution of division times, and are not spatially oriented. Typical of such models is that of Alberghina (Alberghina, 1977), in which a cell divides if its production of ribosomes and proteins reaches a threshold level. The environment can affect the cell by adjustment of reaction rates in the equations modelling the biochemical processes.

While intracellular models of cell division could be used in looking at a population of cells as a whole, it is more likely that statistics generated by such models would be used directly in obtaining a probabalistic model of growth in a population, thus "lumping" the biochemical mechanisms within the cell. Such techniques are employed in Chapter Six.

University of Alberta

3.5 The Experimental Frame

The task of recognizing the experimental frame, through which the real system is to be viewed for modelling and simulation, is that of identifying both the information going into populations of cells in a developing organism and the measurable results which experimenters find in such organisms. Such input and output will be peculiar to the particular system under study, but some general remarks may be made concerning experimental frames for developmental systems of interest.

In order that organised patterns of differentiating cells may be visualized with clarity, an initial decision is

made to study systems which are either inherently two-dimensional or may be viewed in two-dimensional cross-section. A drosophila disc is essentially a two-dimensional structure, while a chick wing's pattern of cartilage and muscle may be regarded in sections perpendicular to the dorsal/ventral axis or to the proximo/distal axis. It should be noted that a base model of development should not reflect this limitation to two dimensions, though lumped models constructed within experimental frames will do so.

The most significant limitation of the experimental frame arises from developmental geneticists' use of "top-down" or difference methods in studying developing systems. A perturbation or sequence of naturally occurring events stimulates division of an initial configuration into subpopulations of cells with mutual differences. Those cellular activities which are common to all cells, and remain common to all cells during the experiment, are not characterised. The experimental frame implicit in such approaches neglects all inputs and characteristics of cells which are not relevant to the developmental changes under study.

In general, inputs to systems consist, then, of "perturbations" (for example, an external chemical signal to a population of cells, or a surgical cut or other disturbance within a population). Inputs to individual cells will be signals thought to be involved in the formation of patterns. Outputs may be of a qualitative

University of Alberta

nature, such as the shapes of developing systems or the normality of an emerging pattern. Further discussion of experimental frames relevant to particular systems will be undertaken in Chapter Six.

University of Alberta

CHAPTER FOUR

The Modelling Framework

University of Alberta

While a particular developmental system of the type discussed in the previous chapter will be characterized by a particular model, there are aspects of such models which are held in common. In particular the base model of the genetic mechanisms operating in cells is almost certainly universal, to the extent that inherited genetic information and environmental factors interact. A particular model for a particular system exists within a modelling framework, sharing some elements with other systems and having some unique characteristics. The modelling framework as a whole may be looked upon as a pool of models, or collection of options, corresponding to various aspects of the developmental process, different developmental systems, and alternative explanations of developmental processes.

4.1 The Underlying Base Model

Consider a cell in a population of cells. It has inherited a state of determination for the activation of portions of its DNA. This state of the DNA is indicative of the cell's present and future chemical activity. The present expression of the DNA will be referred to as the cell's 'functional nature'. Two cells which are in different determinational states (have different activation of DNA) have at least potentially different functional natures.

Consider a population of cells which are mutually undifferentiated. They are in the same state of activation of their DNA, and they have the same functional nature. Now an event occurs which divides the population into two groups with different activation states in their DNA; that is, they are distinct in their determination. It may not be known what the difference between them is, but at least one such difference exists. The difference, whether it be one or a hundred sites of different activation on the DNA, is the result of one determinational event and can be symbolized by one "switch". Such a switch may be likened to a sensor gene in the Britten and Davidson model. The sensor gene acts as a key to the activation of other levels of control over a certain number of cell products.

The situation is pictured in Fig. 4.1. Let the two groups of cells be A and B, and assign switch codes of 1 0 and 0 1 to them respectively. The fact that the codes are

University of Alberta

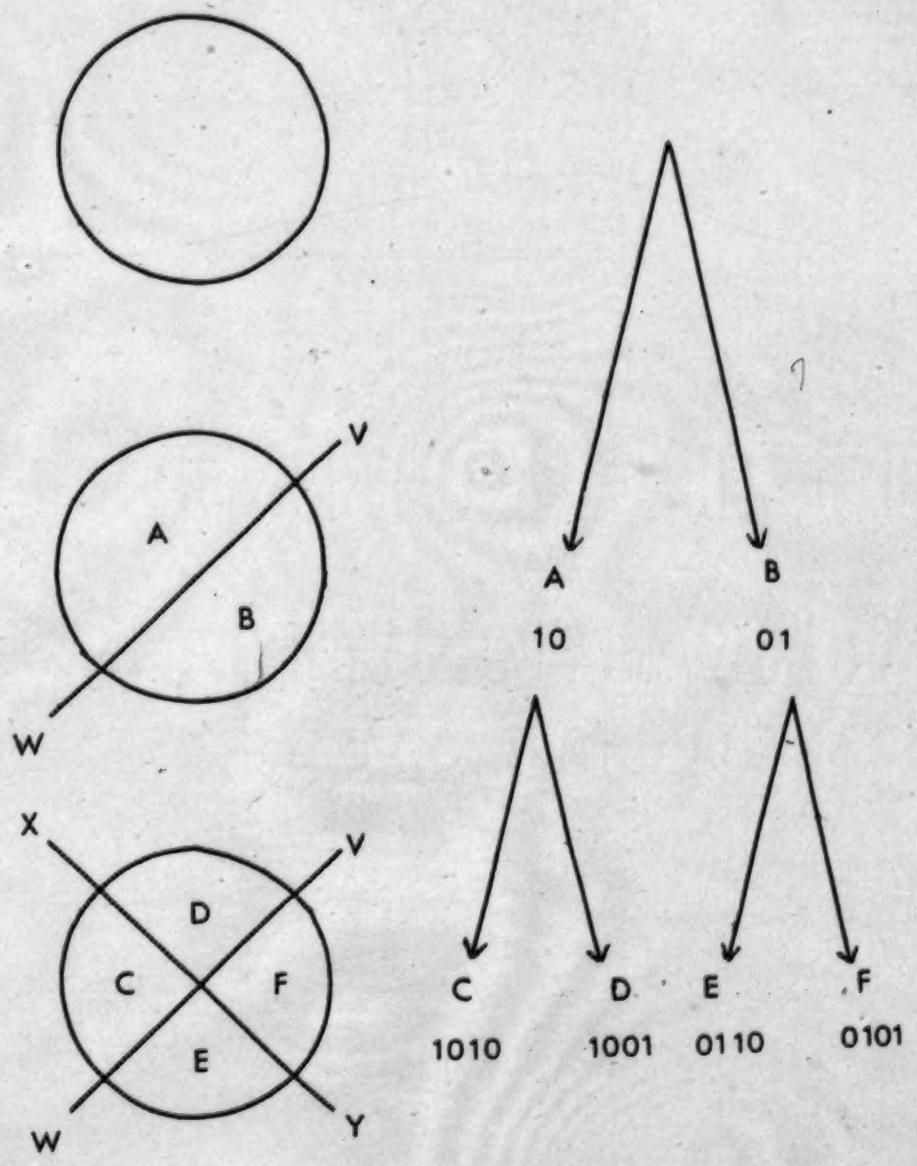


FIGURE 4.1 Coding in Subdividing Populations

different reflects the occurrence of a determinational event. The 1 in each code reflects the activation, or enabling for future activation, of specialized functions for that group of cells. (If the subdivision had been tertiary, codes of 1 0 0, 0 1 0, and 0 0 1 could have been assigned.) The Hamming distance between the codes for A and B cells is 2; this reflects the fact that 2 things would have to take place if A cells were to become B cells: a return to indeterminate state (1 0 to 0 0) and re-determination(0 0 to 0 1). In terms of a representation of determination as a decision tree, the "distance" between two groups of cells at a particular moment in time (at the same level in the tree) may be regarded as the number of branches in the shortest path connecting them.

It may be noted that the codes 1 0 and 0 1 may model equally well a repression of activity as indicated by the 0 in each code. In a repression model, the indeterminate state would be represented by 1 1.

Similarities between this characterization and that of Kauffman (1973) will be noted; however, the code proposed here is intended to have greater significance. Kauffmann assigned his codes at random, the only condition being that Hamming distance should correlate with frequency data. In the proposed model the code for a group of cells accumulates as time goes on and determinational events occur. The code is thus a history of each cell in reference to all other cells since they were relatively indeterminate.

University of Alberta

Furthermore, the code for a group of cells is a mapping onto the DNA, as each 1 represents activated or enabled portions of the DNA.

Assume for example that the different functional nature of the cell groups is to at least some degree expressed in terms of present functional activity. It may not be known what precise chemical activity is exclusive to cells of type A (1 0), but what is known is that all A cells have the same functional nature prior to any subsequent determinational event. That is, all A cells are doing (or can do) the same things, while cells of type B are doing (or can do) at least one thing that is different. It is easy to see that gradients of information can be set up as a result of this differential activity of cells in the population. Consider, for example, diffusion through the entire population of a chemical produced only by cells of type A. A simulation experiment could be set up to see whether the pattern of concentration resulting from such a simple diffusion exhibited an anomaly along the line WV in Fig. 4.1. Thus, without knowing the actual nature of activity of cells of type A, it is possible to study patterns arising from those cells' expression of their nature in the form of signals to the entire population or portions thereof.

If clues do exist as to the functional nature of cells of type A, this information can be used in determining how signals would be transmitted. For example, simple diffusion might be appropriate. If data are available about the

University of Alberta

nature of boundaries between compartments, such information will again be reflected in the way information is allowed to travel. If it is known what shape divisions will arise, and when, different transmission possibilities may be tried in order to throw light on what the actual signals may be. In this way, a model and experimental data can be mutually sensitive.

Each cell in a population has the above code of binary digits which symbolizes its state of determination with respect to the rest of the population. It shares this code with all cells like itself. In addition, however, it has local information in the form of concentrations of chemicals, fields of ionic potential, pressure, and other physical realities in the cell's environment. These quantities may be loosely classified into 3 groups:

(1) external forces, e.g. gravity, organism-wide hormonal flow;

(2) processes common to all cells, e.g. fundamental life support metabolic functions;

(3) local levels of specialized functions; a cell may have its code in common with certain other cells and yet differ from them in the concentration of products activated by the code, perhaps because of positional differences or different timings of the onset of new functional activities.

Using the proposed difference model, as many of the physical influences in the above three groups as are thought to be of interest can be explicitly considered in designing

University of Alberta

a particular modelling experiment; that is, an appropriate experimental frame can be selected. For each of these chosen quantities an account of changes occurring can be kept as a state parameter of the cell. The code of digits will be a key to which of the functions of type (3) a particular cell is pursuing actively (it may be receiving signals from cells with different codes). The concept of each cell's state can be illustrated with two examples.

(1) A population of cells of uniform determination (no code) is subjected to the directed flow of a hormone, and a pressure gradient exists, due to gravity, which influences this flow.

Each cell is characterized by its concentration of the hormone and the local pressure. The value of these as a function of time is given by the solution of an equation of hydrodynamic flow across the population. While other processes are in operation in the cells, these two are the focus of attention. The desire may be to investigate patterns arising in the concentration of hormone.

(2) A disc with two compartments is viewed in terms of the differential functional activity present.

Production of A-specific signals, keyed by the 1 in the code 1 0, starts at the determination event, while B signals, keyed by the 1 in the code 0 1, start simultaneously. The 0's indicate input-only fields for signals. If each cell produces a unit quantity of signal in a unit of time, patterns of relative concentrations of

University of Alberta

signal can be studied. Each cell is characterized by the concentration of the two signals.

A genetic model has been proposed here which views a developmental system as a spectrum of cells, each type of which displays a certain behavior reflecting its determinational state, the pattern of activation in its DNA. This model is the base model for developmental systems. The base model is designed to embody the ways in which the activational state, as symbolized by the code of genetic switches, interacts with the behavior to produce determinational events in an organized sequence. The model is dynamic, flexible in the scope of its attention (the whole organism or any subgroup of cells), and quantitative.

A model proposed by Garcia-Bellido (1975) bears an interesting similarity to the ideas presented above. As a result of his experimental work, he has identified the "homeotic" genes as controlling the development of whole compartments; that is, he has proved the existence of a single gene controlling the action of a battery of producer or "realizator" genes in a manner similar to that envisaged by Britten and Davidson (1969). The homeotic gene, he suggests, is switched on by the product of an "activator" gene in conjunction with an "inducer" molecule. The activator gene is presented as a means of reflecting the genetic inheritance or determined state of the cells, while the inducer molecule is a means of involving local information, thus obtaining positional specificity of the

University of Alberta

homeotic gene. The inducer could be a signal from outside the population of cells, or a signal from cells of a different type. The quantitative model suggested above uses the present activation code to show the determined state of the cells, while local information, as achieved through the functional quantities maintained for each cell, is used to obtain positional specificity of new activation code. In essence, the two models have the same goal, to show the interaction of present DNA activation and local information in producing new DNA activation.

In further support of the approach taken, it may be noted that Lewis (1976) has discussed the discrete nature of compartmentalization events in response to threshold values of continuous gradients of information. He concludes that such development is appropriately modelled in terms of "genetic switching circuits".

Restriction of the base model by the experimental frame results in characterization of cells and cell populations in terms of a difference model; that is, only those portions of the genetic code, and those related activities, which express accumulating differences between cells during a developmental process, are maintained explicitly.

Developmental events occur in an organism because of the recognition of the attainment of critical values of certain behavioral parameters in the cell population. Recognition of such an event occurs when critical values are reached in a particular subcollection of cells, along a

University of Alberta

curve which becomes a compartment boundary, for example.

4.2 Common Elements of Lumped Models

There are two significant simplifications common to all cases of the production of lumped models capable of translation into iterative specifications interpretable by digital computers. First, any continuous time base has to be approximated by a discrete one. The appropriate length of the time step will depend on the particular model being constructed. Second, in the absence of true parallel processing for individual cells, any activities which are in reality simultaneous must be approximated by serial processing. The degree to which this approximation is successful will depend on the treatment of particular processes in particular models. The validity of these simplification steps must be decided in behavioral terms; that is, in terms of the relationship between a particular lumped model, in a particular experimental frame, to the base model and real system.

The fundamental genetic mechanisms in the base model of development translate into structures which are common to all lumped models.

4.2.1 Characterization of Individual Cells

A cell is an entity which has local information of both quantitative and operational kinds. At a given time, the state of the cell is determined by certain local values of

University of Alberta

variables, including global forces such as pressure, metabolic levels, and concentrations or "signal levels" corresponding to expression of the genetic nature of the cell and its neighbours. In a particular modelling situation, a subset of these variables, those directly involved in the developmental process of interest, will be recognized as the state set of the cell.

A cell receives input from its immediate environment in the form of information about the state of those neighbours with which it is in contact. If it is in physical proximity to another cell, but a barrier to certain signal exists between them, no input of these values is transmitted. The cell communicates information about its own state to its neighbours in a similar manner.

A cell contains operational information in the form of its activation code. This information is a set of switches indicating which of its variables change because they form an active part of the cell's expression of its genetic nature. These indicated variables are manipulated by the cell itself and do not merely change passively due to signalling between cells.

University of Alberta

4.2.2 Characterization of a Group of Cells with the Same Genetic Code

A group of cells of identical genetic nature (having the same set of switches) comprises an element in the formation of patterns in the developing organism. The group's component cells, therefore, are considered linked together in a structure which has shape and position in the total population. A "code group", then, is defined by a list of cells of identical genetic state, together with their positions, and shape parameters of the whole group. Description of shape is dependent on the particular system; for example, if the group is a coherent patch, its shape could be described by a list of perimeter cells and their positons.

A code group may itself be regarded as a system with a set of state variables, as a single cell is defined at a lower level in the modelling hierarchy. One element of the state variable set is the string of genetic switches. The genetic activation code changes upon the recognition of a developmental event in the component cells of the group. Another element of the state set is the shape description. A perimeter list will change during growth of the group due to cell division, or because of division of the group into subgroups during an event such as compartmentalization.

Thus changes in the state variables of a group are mediated by events on the cellular level. It is appropriate in most modelling situations to define a code group as

University of Alberta

autonomous (input free) and to consider input only at the cellular level. Output of a code group would depend on the particular model; it might be, for example, the shape description.

4.2.3 Characterization of a Developmental System

A developmental system is a collection of code groups of cells, each group having a distinct genetic nature. The system may be described by listing these groups and locating them relative to each other. The shape description contained within a group may then be used to construct a picture of the developmental system as a whole.

A developmental system changes by the addition of new groups (corresponding to newly specialized cell types) to the list (or, conceivably, by deletion of groups).

University of Alberta

4.2.4 Cell Division

Growth is a process inseparable from development. The formation of patterns of cell types in an organism depends on the mechanism of cell proliferation as well as on events of a genetic nature (Wolpert, 1978). Any spatial configuration must depend on shape and the rate of expansion of the organism. Modelling the growth process is therefore an essential part of modelling any developmental system.

The shape of an organism is controlled by the location, timing and direction of cell divisions. A particular cell's decision to divide is made on the basis of an intracellular

model (for example, a cell cycle clock, or threshold value of a metabolite) or on the basis of an intercellular model (for example, threshold level of a "growth hormone", or a probabilistic model of division). If a cell does decide to divide, it may do so in a preferred direction, for example in the direction of least pressure, or down some other gradient. A pool of growth models containing such options must be made available to aid in the construction of models of particular developmental systems.

4.3 Formal Characterization of Lumped Models

4.3.1 Specification of Individual Cells

A cell is a structure

$C=\langle T,X,W,Q,f\rangle$

where

T is the time base

X is the input value set (range of possible values of signals entering the cell)

W, a subset of (X,T), is the input segment set

Q is the state variable set

f is the state transition function $f:QxW \dashrightarrow Q$

The function f yields the state at the end of each input segment. C is an input-output system if W is closed under composition (successive I/O experiments can be performed) and f has the composition property, that is, for $w(1),w(2) \in W$ and contiguous,

$f(q,w(1)\bullet w(2))=f(f(q,w(1)),w(2))$

University of Alberta

Example:

T is a discrete time base

X = {z(t)|0 < z(t) < Z}, the concentration of substance x entering the cell from its neighbours in time step t

W is a subset of (X,T)

Q is the state variable set, Q=x, the concentration of substance x in the cell

Y is the output value set, {y|y an integer, x-1<y<x}

f is the state transition function, modelling the concentration changes due to reaction and diffusion in the system.

4.3.2 Specification of Code Groups

A code group is a structure

G=<T,Q,Y,f,g> where

T is a time base

Q is the state variable set

Y is the output value set

f is the state transition function, f:Q --> Q

g is the output function g:Q --> Y

Example: Let

T be a discrete time base

Q = (s,L,D) where s is a string of binary digits symbolizing the genetic activation state of the group;

L is the list of cells in the group with their addresses;

D is the list of perimeter cells.

Suppose two kinds of processes are going on in the

University of Alberta

group, cell division and production of a substance x. Suppose that if concentration of x exceeds a critical level c in over 80% of the cells a genetic event will occur changing the value of s from s(1) to s(2). Then f may be given by

f(s(1),L(1),D(1))=(s(2),L(1),D(1)) if the genetic event occurs

=(s(1),L(2),D(2)) otherwise

where L(2) is the list of cells and daughter cells and D(2) is the new perimeter list after cell divisions during time t.

4.3.3 Specification of a Developmental System

A developmental system is a structure

<T,Q,f> where

T is a time base

Q is the state variable set

f is the state transition function f:Q --> Q

Example: Let

T be a discrete time base;

Q = L, a list of code groups with associated location parameters.

Suppose L changes whenever a code group undergoes subdivision into two code groups. Then f(L(1)) = L(2) where L(2) is L(1) minus the dividing group plus the two new subgroups.

University of Alberta

f(L(1))=L(1) if no genetic event occurs in t

=L(2) otherwise.

4.4 Variable Elements For Individual Models

It may be seen that while the structure of the cell, the code group and the developmental system are determined by the modelling framework, the individual elements within the structure are not constrained. In the cell, all the sets T,X,W,Q,Y are free to be defined for a particular developmental model. The transition function f is constrained only by considerations such as a desire that it obey the composition property (if a cell is to be an I/O system) or to be time invariant. Similarly the elements of the code group and organism structures may be defined freely.

The formal specification of these structures cannot be given in iterative form because of the lack of specification of these free elements, in particular of the time base. When experiments are performed on particular models, a formal iterative specification with all free elements fixed will be given.

University of Alberta

CHAPTER FIVE

Implementation

University of Alberta

Corresponding to the modelling framework describing structures and elements common to all the particular models of developing systems to be constructed, there is a computing framework implementing these common features. The framework consists of modules incorporating the separate entities and processes present in the underlying base model described in the previous chapter. Variable elements specifying particular models are introduced either by parameter definition or by selection of optional modules in the "model pool".

5.1 Choice of Programming Environment

Two overriding considerations influenced the choice of programming environment for the project. Firstly, a developing system is a dynamic structure, with a spectrum of

activities in the system which can change with time, and proliferating cells. Of paramount concern, therefore, is the need for convenient (and, hopefully efficient) handling of dynamically changing data structures. Secondly the parallel nature of the model, with different processes going on simultaneously in different cells, influences a decision concerning possible use of available simulation packages.

It has been said that the appropriate time base for the base model is continuous, because although discrete developmental events occur, they arise through cumulative conditions in individual cells. Such events cannot therefore be scheduled in the continuous time base but must be recognized as they occur. The model is, however, of the cell space type, and efficient handling of the representation of parallel processes by sequential means has been accomplished in some such models by use of discrete event strategies (Zeigler, 1976; Hogeweq, 1978). The technique involves processing only those cells which have a possibility of undergoing change at the next time step. Thus a "next event" list is maintained and languages such as SIMULA become useful in minimising processing. SIMULA moreover simulates some aspects of parallel processing for the user.

However the benefits of a discrete event approach to cell space modelling depend on the level of activity in the cells. It becomes efficient only if a significant number of cells does not change at each time step. The developmental

University of Alberta

model has a high level of activity; it is probable that all cells are engaged in some relevant activity all the time.

While packages for continuous time simulations are available, none of these "differential equation" packages is orientated towards cell space models, and the type of data-handling required militates against their use. Typically, activity in a continuous system is modelled by one differential equation, and structures in which activities are locally mediated cannot be treated by the package.

The language chosen for implementation of the lumped model is ALGOLW. It has the right kind of dynamic data acquisition handling. In ALGOLW there are "pointer" data types, in which data is stored and referenced automatically by means of pointers, thus allowing linked allocation of storage. Moreover the data thus referenced can be of mixed type, thus allowing a variety of pertinent information to be maintained for the structure. Pointers, and therefore the data structures to which they refer, can be allocated (and destroyed) dynamically, and thus there is natural expression for structural events such as cell division. ALGOLW has efficient numerical capabilities, with an interface to FORTRAN facilities if desired. A modelling system implemented in this language should be widely understandable and have adequate flexibility.

University of Alberta

5.2 Treatment of the Time Base

The continuous time base of the base model is approximated in the lumped models by a discrete time base. The size of the time step is a parameter determined by the particular model being studied. For a given model, the time step may also be adjusted in response to the generated data. Within one time step various activities may be taking place, for example, cell divisions and changes in several state variable values. In general it will be most efficient to complete all activities for each cell before proceeding to the next cell, as a minimum number of storage operations are thus performed, but the option exists for processing all cells for each activity in turn if such should appear necessary. Further discussion of issues concerning the treatment of time will be discussed in the context of particular models in Chapter Six.

University of Alberta

5.3 Data Structures

The structural elements introduced in the previous chapter (the individual cell, the code group, and the organism) are described by a set of data structures in ALGOLW.

5.3.1 Cells

A particular cell's internal character is defined by the values of its state variables (parameters such as clonal markers, local concentration values, and the genetic code).

In the ALGOLW implementation of a particular model, the appropriate binary code string and parameter fields are held for each cell in a "state" record of a declared structure. Parameters may include local concentration of relevant substances. A state record for a cell is pictured in Figure 5.1. As cells are created initially or through cell division such records can be allocated through invocation of the structure declaration.

When considered as a member of a population, a cell has position and neighbourhood relationships. A two-dimensional population of cells is approximated by a hexagonal grid; that is, cells are pictured as being centred on points arranged hexagonally on a plane. This choice was made for the following reasons. In cells under no distorting forces on a plane, spherical closest packing most nearly approximates nature, in that cells appear in reality to be in physical contact with between 5 and 6 neighbours (Lawrence, 1975); there is no ambiguity regarding neighbourhood relationships due to local symmetry; and arbitrary curves joining cells, as for example along population boundaries, are more easily approximated than on a conventional rectangular array.

The hexagonal grid of cells is defined by a set of "position" records, each containing the hexagonal address, six pointers to similar records for the six adjacent cells, and certain other fields. A position record for a cell is pictured in Figure 5.2. Neighbourhood relationships between

University of Alberta

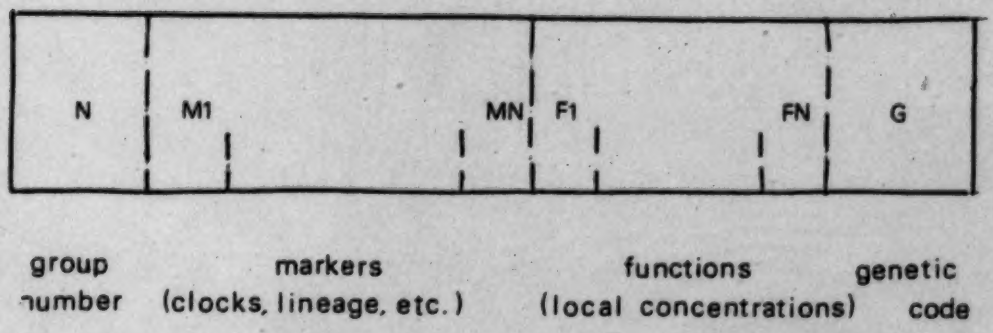


FIGURE 5.1 State record for a cell

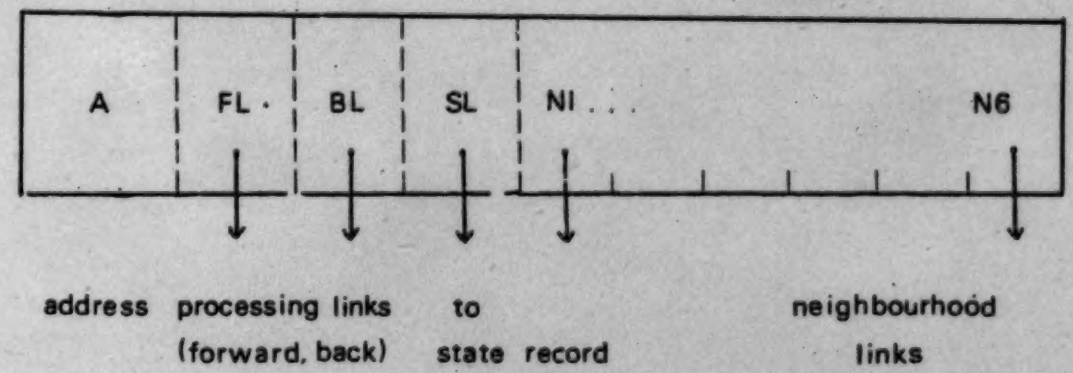


FIGURE 5.2 Position record for a cell

cells are pictured in Figure 5.3. One of these fields is a pointer to a "state" record, symbolizing the location of a particular cell of individual character or "name" at that hexagonal address. It may be noted that if a cell moves, for example in response to cell division in its vicinity, this movement can be accomplished by letting the "position" record for its new position point to its "state" record or name; thus neighbour links do not have to be changed as it assumes a new position, as the neighbour links are associated with positions and not cell names. Other fields in the position record will be pointers associating the position (and, via the pointer to the "state" record, a particular cell) with a processing order and with a code group of genetically identical cells.

It may be noted that in the formal structure for a cell, $\langle T,X,W,Q,Y,f,g \rangle$,
the elements W,Q, and f have been translated into computer terms. W, the set of input segments accessing each cell, is specified by the neighbour links to the cell and by a given starting condition in an array. Q is the set of state variables held in the cell's "state" record; and f is defined by the genetic code's switching pattern on the state variables. Furthermore structures exist for selecting Y and g in a given experiment; for example, the output function g may select of cells of identical genetic nature, in which case the list of cells linked together by the group link can be output in linear or graphic form.

University of Alberta

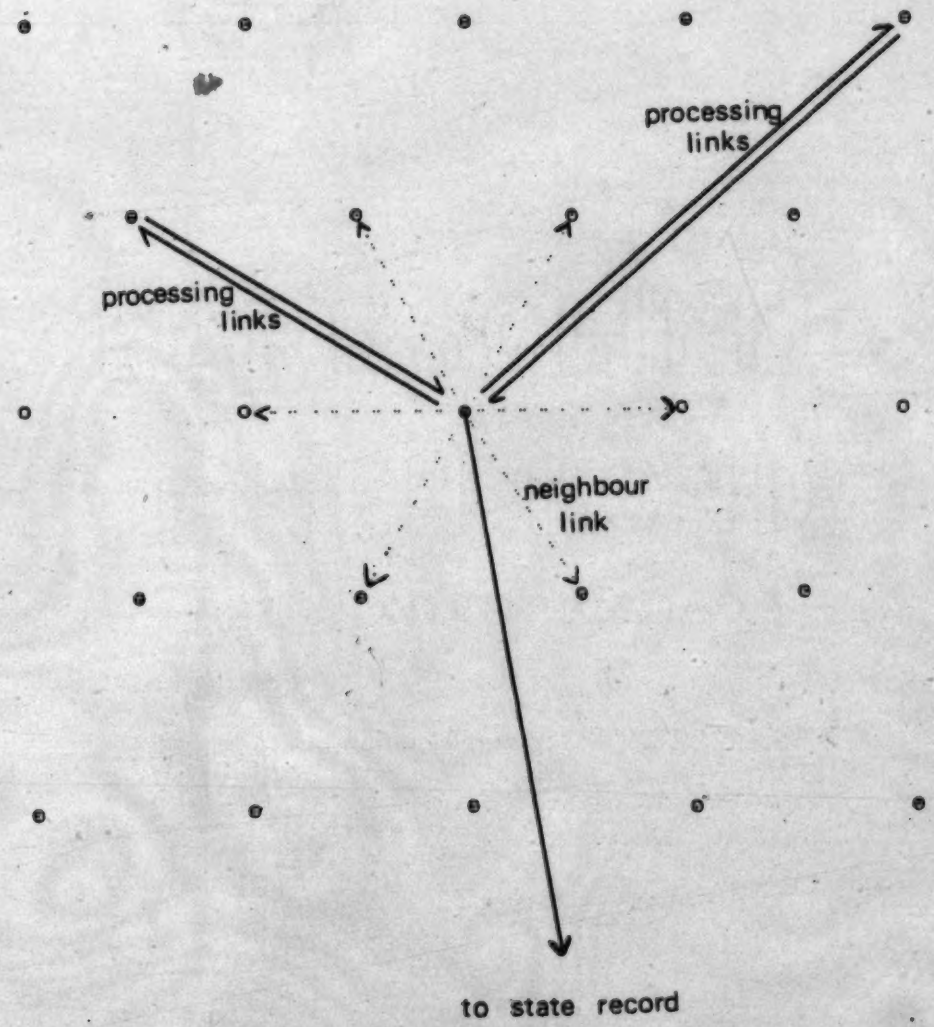


FIGURE 5.3 Neighbourhood relationships between cells

Positions in the hexagonal array are organised into processing lists and code group lists. In order to locate a cell by its position, unless further information is given, no alternative to searching along one of the lists exists. Such searches are expensive because of the number of storage references involved in following the pointers. Whenever this kind of problem arises, local "hash tables" based on position coordinates can be set up; that is, cells in a list can be located in a linear array by a function operating on their addresses. The entry in each position of the hash table is a pointer to the actual cell located there (the machine address of the record associated with the cell). When no longer needed these tables are destroyed.

5.3.2 Groups of Cells of Identical Genetic Nature

The formal specification of a code group has been given as <T,Q,Y,f,g>.

The state variable set consists of the genetic code common to cells in the group, a list of member cells, and a shape description. The list of member cells is given by following the "group links" in the position records. This list may therefore be specified by a pointer to the first cell in the linked list. The genetic code of the group is carried by each member cell. A shape description such as a perimeter list can be constructed for a given modelling experiment. For example, a "shape" record class may take the form of a pointer to the "state" record of the perimeter

University of Alberta

cell and a pointer to the next "shape" record. Then the shape list is given by a pointer to its first record. A "shape" record list is pictured in Figure 5.4.

The transition function f operating on the state of a code group must be specified for individual developmental models in the form of a routine which implements a growth model or recognizes critical configurations in the member cells. The output set Y and function g must also be specified for particular requirements, but advantage can be taken of the existing structures (for example if Y is the perimeter list, g merely lists it by following links).

University of Alberta

5.3.3 The Developmental System

The developmental system has been specified formally by <T,Q,f>, Q being a list of code groups with associated location parameters. The list is defined by a record class known as "groups" which gives each group an identifying number and associates with the number pointers to the member list and shape description and appropriate location parameters. The transition function must again be specified for individual developmental models. The "groups" record list for a developmental system is pictured in Figure 5.5.

5.4 Sequential and Parallel Processing

The base model is parallel in two senses: different cells are undergoing changes simultaneously, and different processes are occurring at the same time. Both kinds of

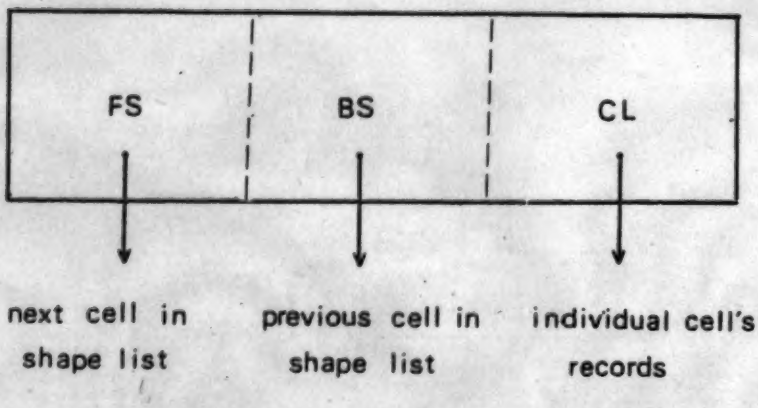


FIGURE 5.4 Shape record list for a code group

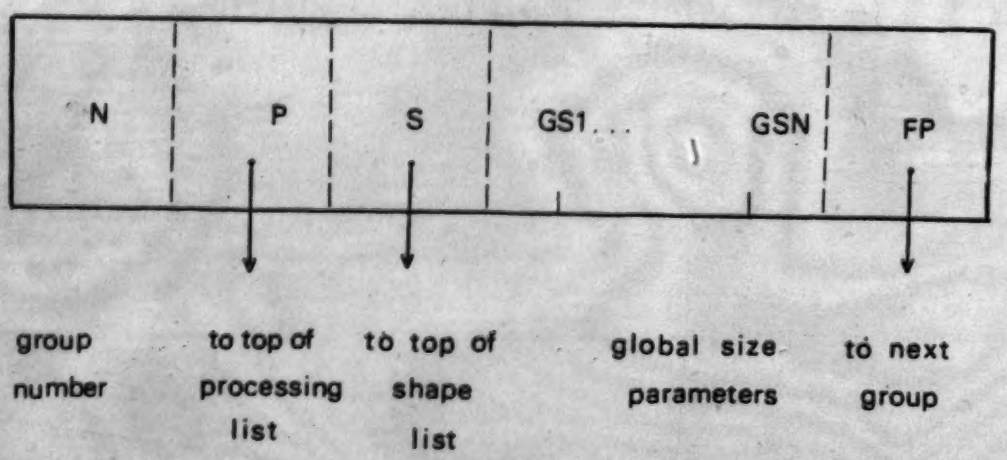


FIGURE 5.5 Groups record list for a developmental system

parallelism must be approximated by sequential processing in the digital computer.

Parallel processes may easily be approximated in a sequential manner (that is, one process completed for a time step, then another) as long as they are independent. Consideration must be given, however, to situations in which processes are not independent. For example two substances X and Y may be produced according to a coupled pair of production functions as they catalyze or inhibit each other. Then the action of the transition function on the state variables symbolizing X and Y must be implemented so as to reflect the interdependence of their action.

Simultaneous activity in the array of cells must be approximated sequentially with caution. For each type of activity, care must be taken that systematic distortions are not introduced, for example by the order in which cells are processed. Models of a particular activity must be assessed with respect to their success in preventing such distortion. In addition extra storage will be needed in order that the "simultaneity" be preserved between communicating cells. If the next state of cell B depends on the present state of cell A, cell A's computed new state must be stored until cell B's new state is computed, and then cell A can be updated.

Problems of parallel to sequential conversion will be considered in more detail in the discussion of individual models in Chapter Six.

University of Alberta

5.5 Representation of Growth

As has been noted, growth is an activity common to all developmental models; in addition, consideration of the implementation of growth models provides an example of processing using the described data structures. The mechanics of cell division in these data structures are is pictured in Figure 5.6.

Cell division and the control of shape in populations of cells has been modelled in terms of two decisions, at each of which various model options can be brought into play. Firstly a cell decides whether it is going to divide or not (decision point A). It does so according to some growth model (whether an internal clock, internal threshold value, external signal, or probabilistic function value). If it does not divide, processing passes to another cell.

The shape of the population may be governed entirely by this first decision, for example if only those cells in a certain "growth zone" are permitted to divide. It is possible, however, that the cell may decide in which direction to divide (decision point B), according to some other model (in the direction of least pressure, or down a gradient, for example).

Growth Algorithm

(1) Set pointer P to head of processing list

(2) Is P null? If yes, stop.

(3) Decision point A: does cell (B) divide? If yes, invoke division model. If no, go to 9.

University of Alberta

(4) Decision point B: In what direction does it divide? Chose random direction, or invoke direction model, setting direction i, $1 \leq i \leq 6$.

(5) Proceed in "position" data structure along neighbour link i until first empty space is reached (link is null).

(6) Create new "position" record for empty space and link to appropriate neighbours.

(7) Move "state" record pointer, processing links and group links outwards along direction i to symbolize expansion of population making room for new cell.

(8) Create new "state" record for new cell, and attach to adjacent position i. Put it before cell (P) on processing list (it should not divide in this time step) and link it to the same code group as cell (P).

(9) Set P to the link of P (go to next cell on processing list). Go to 2.

University of Alberta

5.6 Problems of Economization

The data structure necessary to implement a model in which cells carry internal information, and in which processing is efficient, is unavoidably extensive. It is likely that in an experiment of some size the space requirements will become too large. There are three ways in which economization of space could be realized when necessary.

The first technique involves adjusting the focus of

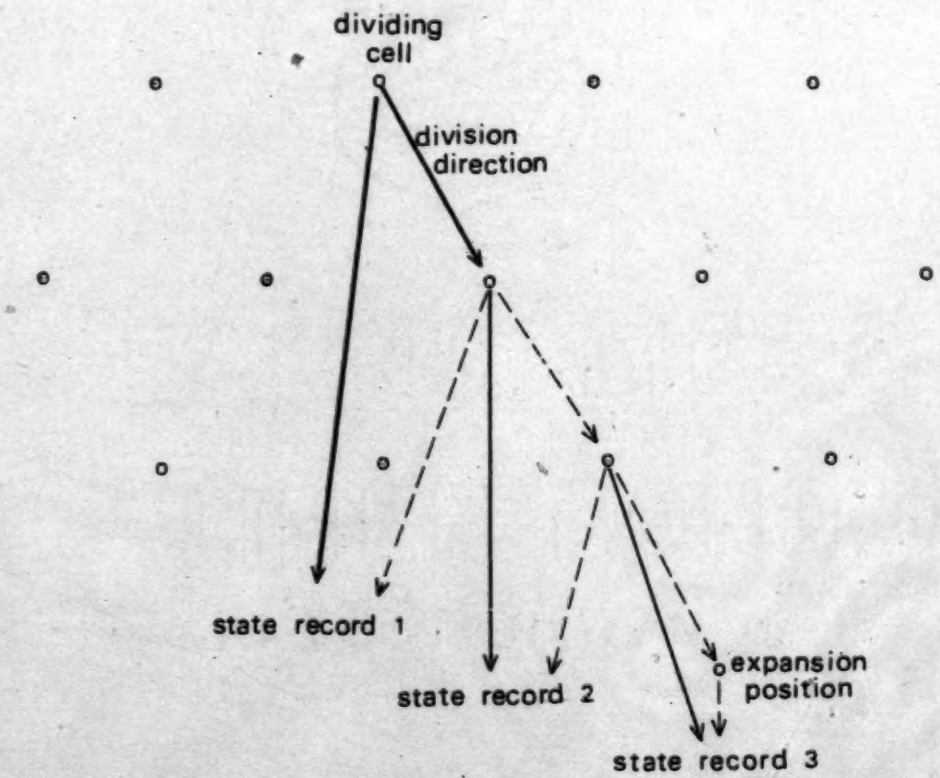


FIGURE 5.6 Division of a cell

attention. For example during simulation of a disc undergoing compartmentalization, different code groups will be formed corresponding to the different compartments. If the disc were to become too large to handle in entirety, the simulation could be stopped and restarted with focus on one compartment. In this case the entire processing list would be replaced by a code group list.

The second technique involves representing a number of cells by one cell with representative characteristics. A mapping of the neighbourhood geometry from the simple case to the averaging case would have to be preserved, and the method would not be applicable in cases where cells in different code groups were interspersed.

The third technique involves a redefinition of the data structures used in implementing the lumped model. As they have been defined, the data structures mirror the structures defined for the model specification in Chapter 4. However a good deal of space, and some time spent in storage references, could be saved by making greater use of hash tables.

Before a processing step, comprising cell division, communication and possible genetic events, the size of the organism is known by parameters such as the extent parameters given for each code group. An estimate can be made of the probable size the organism will attain during this step and a hash table set up to cover it. This hash table would contain, in each row corresponding to each

University of Alberta

position, all of the structures as previously defined except the "state" record. It would contain fields for the processing link (row number of the next cell to be processed), group link (row number of another cell in the same group), "state" record link (pointer to the record) and perimeter link for shape description (null for internal cell, row number of next perimeter cell for external cell). Neighbourhood links would be obtainable by arithmetic on the row numbers of the hash table.

The hash table representation is more economical but does not mirror model structures as closely as the record classes defined previously. Processing could be conceivably less convenient. Until the modelling framework has undergone the first developmental phase, the previously defined data structures will be maintained.

University of Alberta

5.7 Validity of the Implementation

Specification morphisms between the computer implementation and the lumped models are determined by inspection of the state transition function. They cannot be discussed, therefore, until specific developmental models with specific transition functions are defined. Care has been taken, however, to ensure a degree of structural faithfulness in defining data structures to mirror the structural elements of the underlying lumped model, namely the cell, the code group, and the developmental system.

CHAPTER SIX

Experiments and Results

University of Alberta

This chapter describes the construction and use of a modelling structure containing many options and strategies for simulating development. In the context of this structure, an "experiment" is the simulation of a particular model in a suitable experimental frame. The establishment of the entire structure, experiment by experiment, can in no sense be an exhaustive process; that is, there are always further options to be explored in modelling a given phenomenon. Attention is given to models of current genetic interest, for example, reaction-diffusion systems; strategies for implementing other models are mentioned where relevant.

6.1 Experimental Sequence

As previously discussed, development involves the

interaction of growth (cell divisions in a population of cells) and pattern formation. In the initial stages of the project, attention was focussed on experiments with models of growth. Subsequently attention shifted to investigation of mechanisms of communication between cells and the formation of patterns of information in groups of cells. Experiments were then aimed at modelling growth and intercellular communication simultaneously. Finally, various models for actual developmental systems were studied.

University of Alberta

6.2 Modelling Growth

Growth may be defined informally as the process by which a developmental system is enlarged by division of individual cells. The process is controlled at both cellular and system-wide levels. Each cell appears to divide in response to attainment of a critical internal level of protein content (Alberghina, 1978). Cell division rate in a system appears to be constant in a given environment and to be responsive to changes in environment (Kiefer, 1968; Wheldon, 1973). Furthermore, the kinetics of cell division may alter, both in response to cell differentiation and in response to attainment of maximum allowable size of the system (Alberghina, 1975, 1977; Kim et al., 1973). It may be remarked therefore that models of growth should embody various levels of control and feedback.

6.2.1 Formal Specification of Growth Models

A growth model is a structure G given by

<Q,T,y>, where

Q is the single state variable, consisting of a linked list of cells defined by their addresses, links to neighbours, and pointers to individual parameters;

T is a continuous time base;

y is the state transition function

y:Q --> Q.

The state transition function y, which maps a list of cells into another list containing new daughter cells, operates in the continuous time base T. Its action is to be represented in discrete time, and in iterative form, in a lumped model given by systems S1 and S2 at the cellular level, and G1 at the group level.

S1 is a system <Q1,I,T1,f1,g1,Y1>, where

Q1 is the state variable set, each element of which consists of the address of a cell;

I is the set of all inputs to S1 (factors in the cell's environment affecting division);

T1 is a discrete time base, with a constant step size measuring the time to consider one cell (cell processing time);

f1 is the state transition function, which replaces one cell by the next on the processing list

f1:Q1 --> Q1 is given by

f1(c)=link(c) , c ∈ Q1. If c=null, system terminates.

University of Alberta

g1 is the output function, discussed below;

Y1 is the output set, consisting of all couplets (a,b),

a $\in$ {0,1}, b $\in$ {1,2,3,4,5,6}.

The output function g1:Q1 x I --> Y, calculates a condition local to the cell in question and specific to a particular model. It returns a=0 if the cell is not to divide, and a=1 if it is. Furthermore it returns the digit b signifying the direction of cell division, which is again specific to a particular model.

The output couplets of S1 form input to subsystem S2, which is a system

<Q2,I2,T1,f2,g2,Y2>.

The system S2 is thus a means of locating dividing cells on a processing list.

where

Q2 is the single state variable set, with domain the positive integers;

I2 is the input set, generated as Y1 for S1;

T1 is the discrete time base of system s1;

f2 is the state transition function

f2:Q2 x I2 --> Q2 where f2(q,a)=q+1;

g2 is the output function; it generates a 1 if a dividing cell is reached in the processing list, 0 otherwise:

g2:Q2 x I2 --> Y2 where

g2(q,a)=0, a=0

g2(q,a)=1, a=1;

Y2 is the set (0,1).

University of Alberta

The system G1 is given by

<Q,IG,XG,T1,fG>, where

Q is the single state variable, consisting of a linked list of cells with addresses and neighbour links;

IG is the input set, a triplet (c,d,e), where

c is the output of S2,

d is the state variable value Q2 of S2,

e is b from system S1 (direction of division);

XG is the set of actual values received from S2;

T1 is the discrete time base;

fG is the state transition function

fG:Q x IG --> Q where

fG(q,0,d,e)=q, q $\in$ Q

fG(q,1,d,e) results in the action

(a) copy entry d+1 in the list q and insert it in position d+2 in the list (set pointers to and from it);

(b) change neighbour links where necessary;

(c) change parameter pointers along direction e effecting cell movements.

The operation of the growth model may be summarized verbally as follows. A structure of cells, specified by a linked processing list with neighbour links, undergoes changes which are mediated dynamically at the level of the individual cell. At discrete intervals of time the structure changes by the addition of a single cell and corresponding changes in position in the surrounding structure. Growth in a given time interval is represented

University of Alberta

by the sum of such individual events arising dynamically during that time.

Implicit in the formal presentation of the growth model are concepts developed by Zeigler(1976) in his Twelve Postulates for valid modelling and simulation. In particular, the "real system" is the set of all cellular structures in nature undergoing cell divisions. It is observable through data such as initial and final shape, number of cells, and clonal distributions (marked cell lineages). The "base model" is the structure G above. The set of "experimental frames" in which G could operate includes all cell structures; however, G is studied here in the universe of two-dimensional structures or organisms observable in two-dimensional cross-section. A particular experimental frame is given by a set of allowable input values (a subset of the environmental influences of set I of S1) and a set of observable outputs (a subset of the observables mentioned above in the "real system"). The "base model in the experimental frame", G/E, is the structure G with Q restricted to those states for which the observables fall within the output set of the experimental frame, an output set consisting of the output set of the experimental frame plus a null output for nonvalid (in E) input segments, the state transition function f restricted to final states which have observables in the output set of E, and an output function yielding these observables for allowable states.

University of Alberta

The operation of the base model in E (a particular experiment) is known to the observer through observations made on the real system. The real system itself is accessible through the totality of observations in all experimental frames.

In the process of modelling and simulating growth, two iterative specifications are developed. Firstly, there is the lumped model (systems S1, S2 and G1), and, secondly, the simulating program. There are two tests of validity of the process which must be made. If the data generated by the lumped model in experimental frame E are the same as the data expected for the base model in E, then the lumped model is said to be valid for the real system in E (the action of S1, S2 and G1 on allowable environmental inputs and initial structures produces realistic final structures for this experimental frame). If there is a specification morphism between the simulating program and the lumped model, then the program is said to be a valid simulator of the lumped model.

The success of the iterative lumped model in portraying reality can only be judged, therefore, by comparing the iteratively generated results with real data. The success of the simulating program in translating the intentions of the lumped model can be ensured by requiring that (1) the input segments of the computer program can all be "translated" into allowable input segments of the lumped model (thus ensuring that extra influences are not at work

University of Alberta

in the simulator), (2) there is an "onto" map from a subset of the state set of the lumped model to the state set of the simulator (every state of the simulator represents a valid state in the lumped model) and (3) any output of the lumped model can be identified with some output of the simulator (ensuring that the simulator has a broad enough scope). Then if both the state transitions and outputs are preserved under translation, a specification morphism holds.

In terms of the modelling of growth, the computer program is designed, for each of the options developed below, so as to accept allowable inputs and initial states for the given experimental frame. The actual state values which may be generated by the program (structures of cells) are restricted to structures realizable in the lumped model (for example, by restricting the structures to the plane). Requirement (3), that any output of the lumped model can be achieved by the simulator, is harder to ensure. It can be inferred only after a multiplicity of simulations has been run. Preservation of the state transition function means that when the representation of a state of the lumped model in the simulator undergoes transition, the resultant state of the simulator represents that particular state of the lumped model under the same transition; that is, a structure generated by the simulator representation of the transition function itself represents that structure which would have been generated by the transition function of the lumped model. That this condition holds must be established by

University of Alberta

simulation experiments (see Sections 6.2.3-6.2.5 below). Similarly, output preservation is established by running simulations.

The major concern in establishing the validity of the lumped (iterative) model as a representation of the real system lies in the replacement of a continuous time base by a discrete one. Thus Section 6.2.2 is concerned with experiments designed to validate this representation.

Before particular experiments in growth are described, some general introductory remarks may be made. Firstly, the model of growth is necessarily complex and multilevelled because it deals with dynamically changing structure, one of the essential challenges in the project, and because it must contain opportunities for different levels of control. Secondly, the system G is represented as being autonomous (input-free). Environmental influences are modelled at the level of the individual cell in system S1, reflecting the fundamental assumption that the processes of development are mediated by events local to cells. Thirdly, it may be noted that several different time bases are employed in the model. The relationship between these time bases should be explored.

University of Alberta

6.2.2 Sequential and Parallel Processing: Experiments with Time

The system G proceeds from one value of its time base T to the next, producing a new structure of cells and daughter

cells. In nature, individual cell divisions, once initiated appear to proceed independently without regard to other divisions. That is, in the context of a single interval of time of base T, all cells that divide do so simultaneously. System G could therefore be modelled with structural validity by means of parallel processing. However systems S1, S2 and G1, which are iterative specifications suited to a digital computer, treat the cells sequentially (as they appear on a processing list). An hypothesis concerning the validity of such processing is presented and tested at various levels.

University of Alberta

Processing Hypothesis: The system G may be validly modelled by sequentially processing S1, followed by S2 and G1.

The processing hypothesis may be tested at two levels. Firstly, the effects of separating the decision to divide (S1) from division itself (S2 and G1) should be investigated. Secondly, any distortions in structure resulting from cells dividing in sequence (in S2 and G1) should be noted.

Separation of S1 from (S2 and G1) is desirable computationally because cells already processed in S2 and G1 are thus incapacitated from dividing in this time step in the base T1. If they were allowed to decide to divide while the division process is in train, the list would have to be re-scanned to accomplish their division, and the process

would repeat itself indefinitely. The question is whether this limitation of the model constitutes a serious difference from the natural situation being modelled. In nature, during an interval of time (which is continuous), cells initiate the division process at any time. Thereafter, however, until the process is complete division cannot again be initiated. Thus if the discrete time step in system G is not allowed to be greater than the cell cycle time, no cells "flagged" to divide in that time step will do so again. However cells not initially flagged to divide may subsequently decide to do so because of a change in local conditions (modelled in S1 by the action of g1 in response to the state of the cell itself and inputs from its environment.) If any discrete system modelling such change in local conditions, is not allowed to have a time step incremented during a time step of G (that is, the time step of that system is an integral multiple of the time step of G), such a change cannot happen.

Thus there are conditions on the time base of G which allow it to be modelled by separation of S1 from S2 and G1. In terms of the relationship of the lumped model of G to a real system it models, the conditions under which a discrete time base can represent continuous time are beginning to be established: the time step must be smaller than cell cycle time, and "small enough" to be sensitive to other processes influencing cell growth.

The above discussion of separating sequential

University of Alberta

processing into S1 and (S2 and G1) has introduced the question of changes in local conditions which are effected by systems other than the growth model. The question cannot therefore be pursued experimentally without the development of such systems. Restrictions on the time base have been established, and attention is now turned to the second level at which the processing hypothesis can be tested. An experiment designed to test any distortion of structure by sequential division of cells is presented.

Processing Experiment: To examine distortions of structure arising from sequential processing.

Experimental frame: Planar structures are constructed using the growth model in various forms. The initial structures are identical but processing is undertaken in varied ordering. Autonomous models are used (growth unaffected by environment, no input), and output is to consist of a mapping of the structures at each growth time step, where cells bear "clonal" markers - that is, the lineage of each cell is given.

Model features: The output function g1 of S1 (by which cells are flagged to divide) gives 1 for all cells (unconditional, synchronous division). Its second component, the division direction, is "uniform random" (generated by a pseudo-random number generator) or in the direction of "least pressure"

University of Alberta

(generated by looking along the six neighbouring directions and finding the nearest empty space).

The transition function f2 of S2, which follows the processing list and replaces one cell by the next, is varied by means of manipulating the processing list. This list can be, in turn, left in its natural order as daughter cells are added, or "randomized" at each growth time step, or sorted consistently into the same geometrically-oriented order.

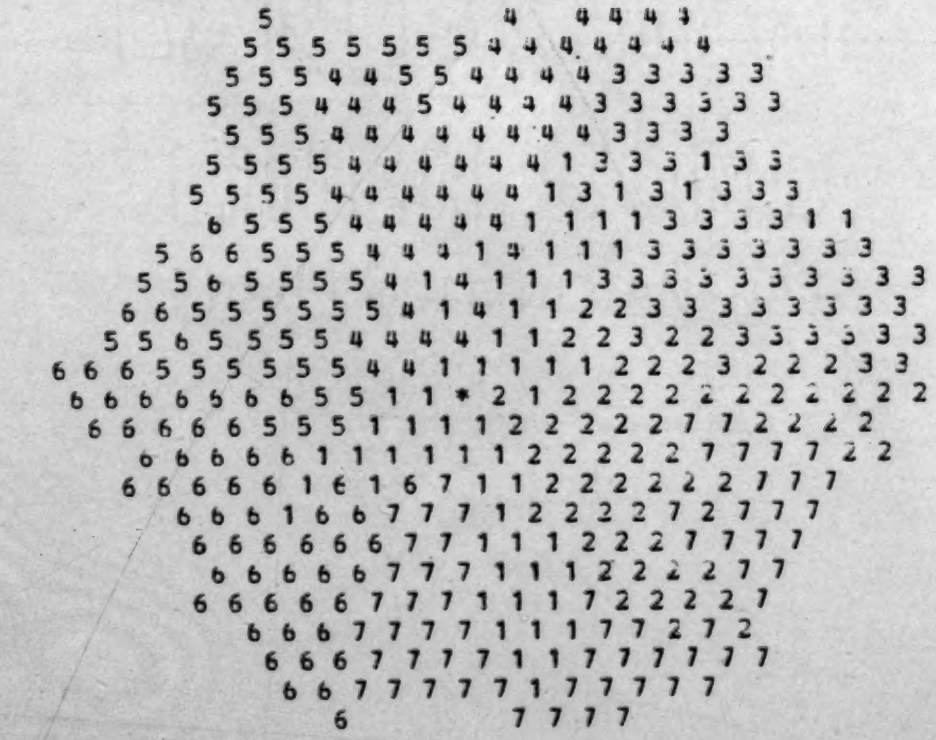
Test Cases: Initial configurations of various types are tested. In particular, a structure which, under the "least pressure" option, may be expected to develop into a structure with coherent clones, is chosen; it consists of six cells surrounding a seventh in hexagonal symmetry. Various irregular structures are studied; under the random direction option, distortion might be suspected if cells of the same lineage form unnaturally large coherent clumps. In all, about 20 structures thought to be of interest are taken through about 8 division cycles.

Results: The output from the hexagonally symmetric initial structure is presented in Figures 6.1-6.3, as it is representative of all results obtained. In no case are completely coherent clones obtained, but there is no striking, systematic difference between results obtained with unsorted, randomized, or geometrically ordered processing. (It should be noted that complete coherence

University of Alberta

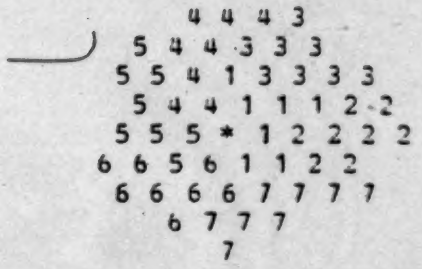
GROWTH CYCLE 0

GROWTH CYCLE 4

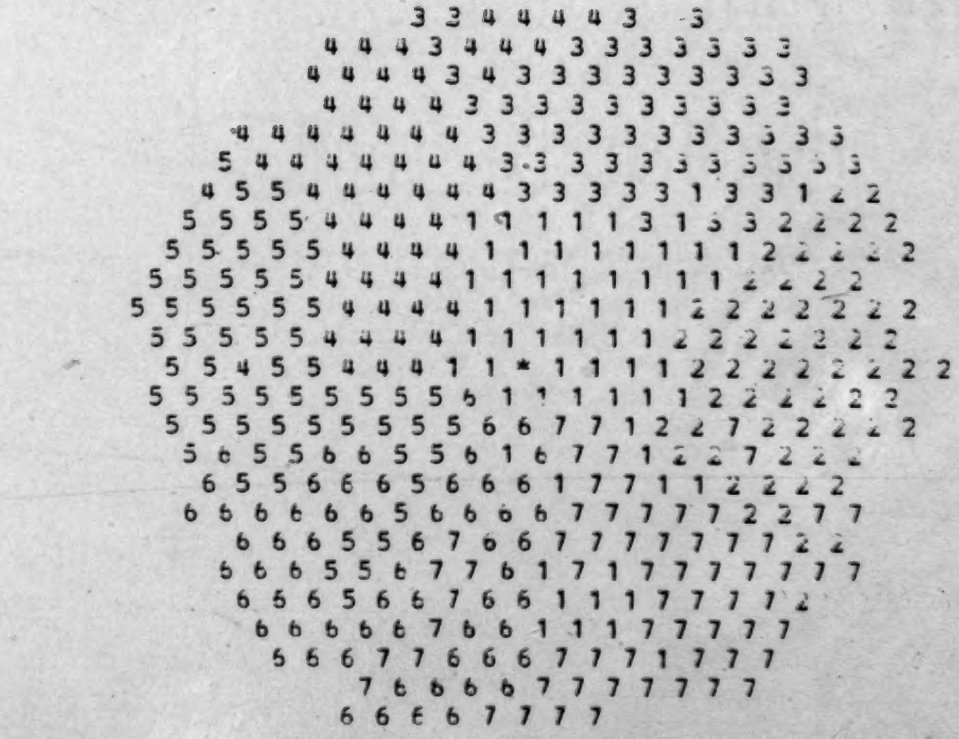


GROWTH CYCLE 7

FIGURE 6.1 Growth with randomized processing list

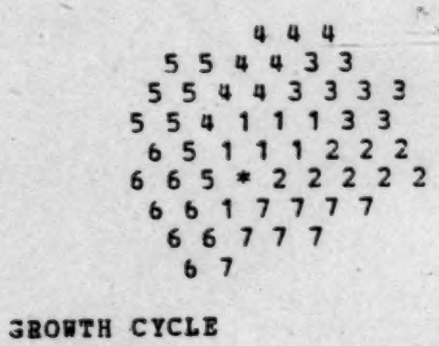


GROWTH CYCLE 4

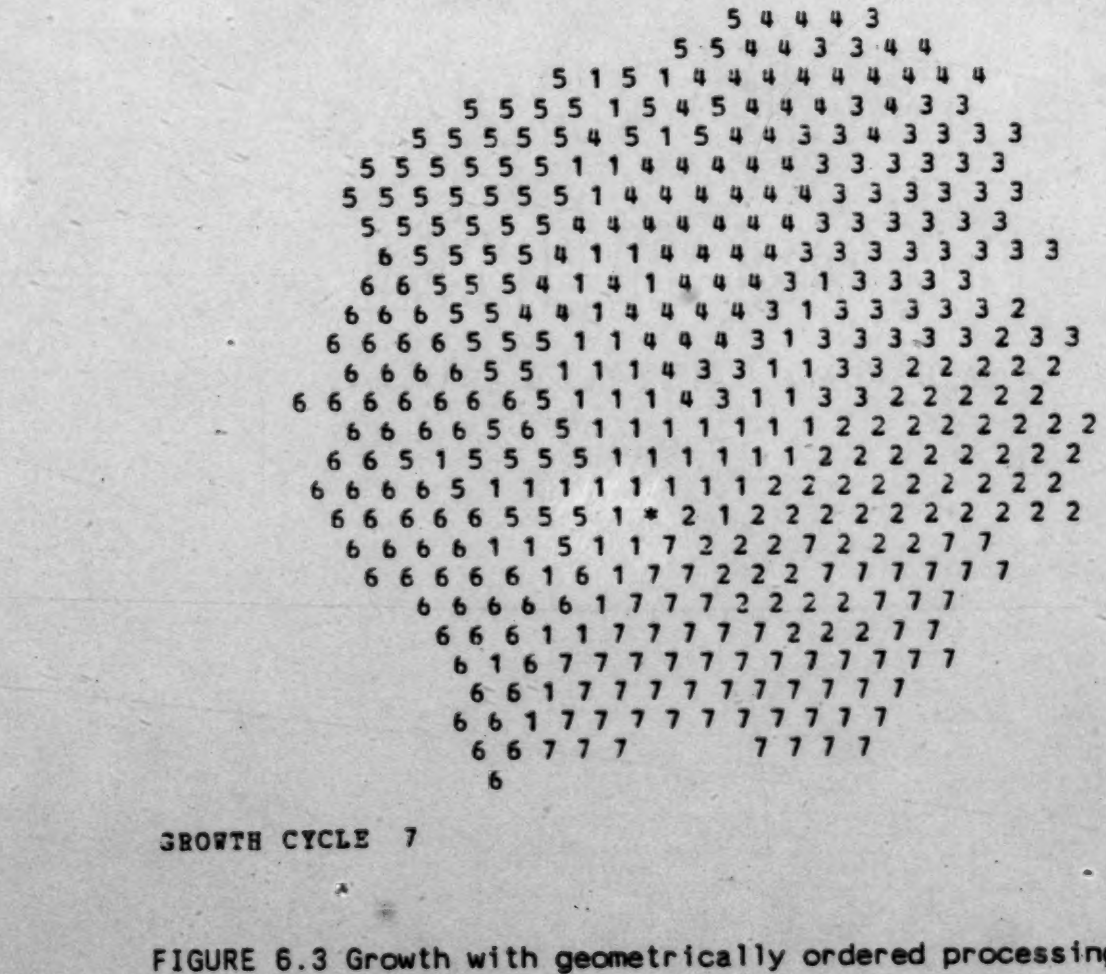


GROWTH CYCLE 7

FIGURE 6.2 Growth with unsorted processing list



GROWTH CYCLE



GROWTH CYCLE 7

FIGURE 6.3 Growth with geometrically ordered processing list

would not be expected even with parallel processing in a hexagonally symmetric situation as there is symmetry in the choice of division direction along the axes of symmetry). Most strikingly, persistent geometric ordering does not appear to result in a distortion in this case. It might be expected that in Figure 6.3, where processing always proceeds from the bottom left, the cell lineages 5 and 6 would be more fragmented due to cell movements caused by later cell divisions; such is not the case.

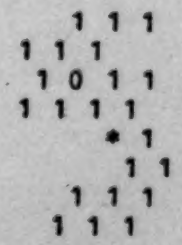
Whatever the processing order, cells dividing towards the direction of least pressure tend to shift irregular initial configurations towards similar convex ones (Figure. 6.4). In order to emphasize this result, the perimeter of the structure is displayed.

University of Alberta

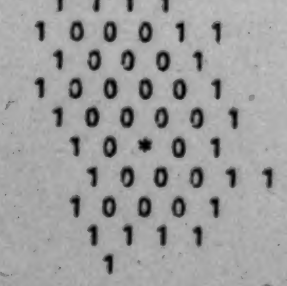
Conclusions: The processing order in the sequential modelling of growth does not appear to distort resulting structures. (The option of randomizing the processing list is retained in case systematic distortion is ever suspected). Based on the results of the processing experiment, and under the above restrictions on the time base, the processing hypothesis is considered established.

Two conjectures may be advanced on the strengths of the results of the processing experiment. Firstly, it may be that distortions do not appear because they are "damped" by other influences which model "dissipative" forces (for

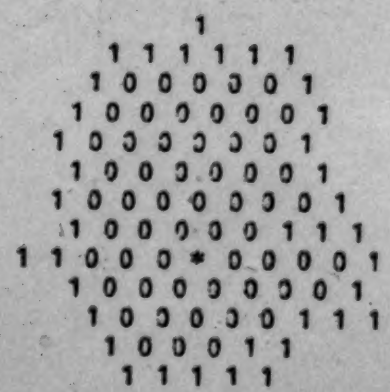
GROWTH CYCLE



1 GROWTH CYCLE 2



GROWTH CYCLE



3 GROWTH CYCLE 4

FIGURE 6.4 Growth from irregularity to convexity

example, division in the direction of least pressure, which brings about a minimum energy situation). If so, any non-uniform ordering of cells will have to be caused by "disruptive" or "reactive" influences and modelled accordingly. This conjecture appears to be in line with events in the real world where achieving nonuniform structure implies expending energy. (Care must be taken if such disruptive influences are modelled, that distortions from sequential processing do not then appear). Secondly, it may be suggested that in modelling systems in which coherent clones appear, processes preserving coherence beyond the state of initial contiguity must be included. These might include imposing a "bond" between cells in the same clone lessening the likelihood of their being separated by other cell divisions. Such bonds could be modelled by recognizing neighbouring cells of the same clone and preferring cell movements which do not separate them.

University of Alberta

6.2.3 Growth of Competing Structures

When the scope of a modelling experiment is such that different groups of cells are processed separately, conflicts may arise between groups of cells growing in physical proximity. An example is compartments growing in a Drosophila disc. Various options have been implemented for dealing with structures competing for space to expand.

The conflict is recognized when a cell selects a division direction and attempts to move cells outwards along

that direction. In searching along the "ray" for an empty space, it finds instead a cell belonging to another group. Subsequently one of the following options may be selected:

Option 1: The second group is regarded as penetrable and cell movement outward along the ray is continued to the opposite border of the second group.

Option 2: The second group is regarded as impenetrable and as exerting a force opposed to growth in that direction. In this case, the direction of cell division is reversed.

Option 3: The second group is regarded as impenetrable, but the cell divides towards it causing a movement of cells sideways along the boundary to accomodate the new cell.

These options have been implemented but not extensively tested, as they do not come within the mainstream of development of the project. Figures 6.5 and 6.6 illustrate the effect of a barrier, either of other cells or of a physical nature, on a dividing population under options 2 and 3 respectively.

University of Alberta

6.2.4 Modelling the Division Decision

The output function g1 of S1 generates a 1 for each cell that is to divide. It does so by some calculation on local conditions (the values of state parameters and inputs). The nature of this calculation is dependent on the particular model of growth selected. Autonomous, unconditional division has been presented in Experiment A. Other options are presented here. In what follows "random"

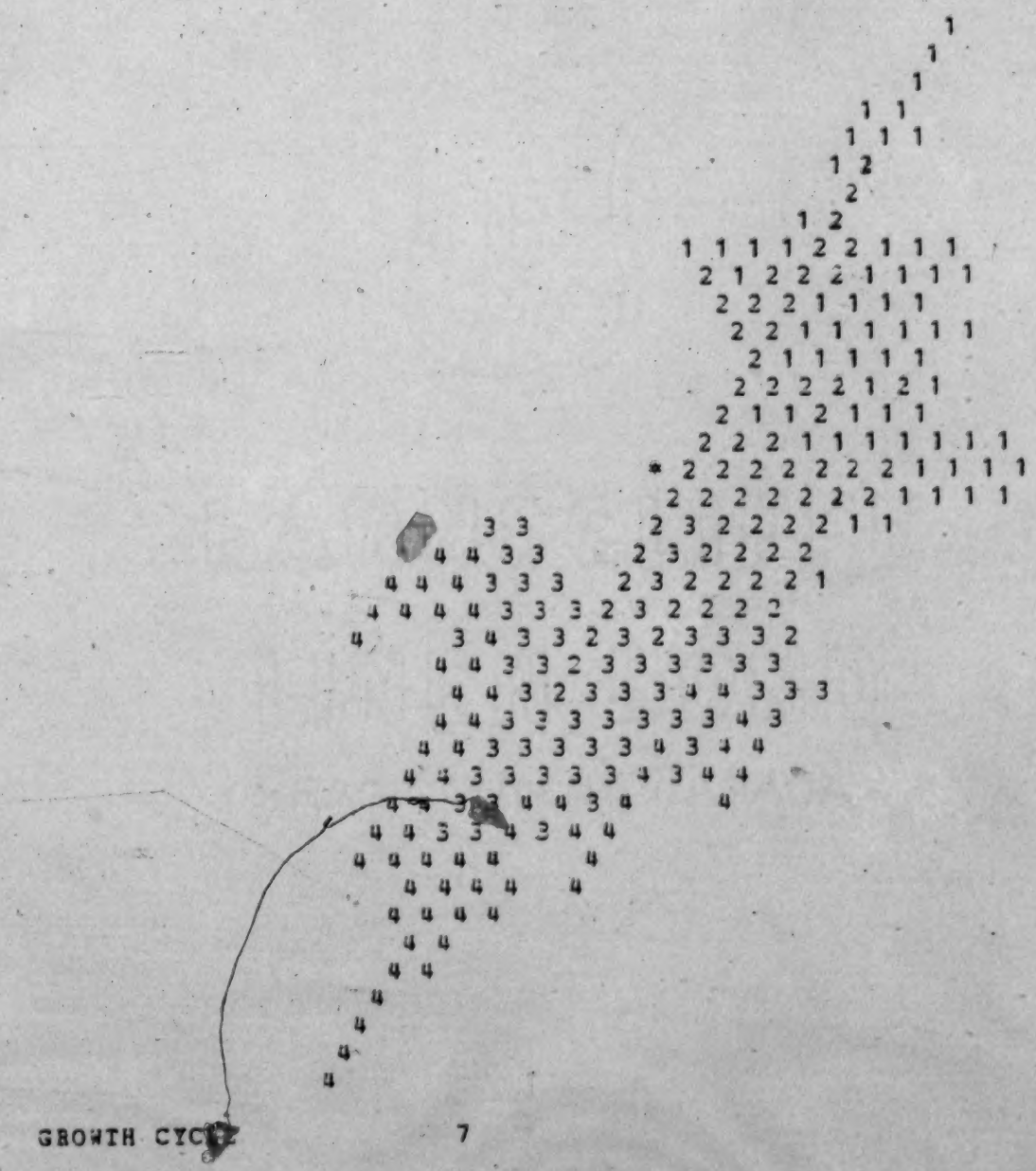


FIGURE 6.5 Growth near an impenetrable barrier

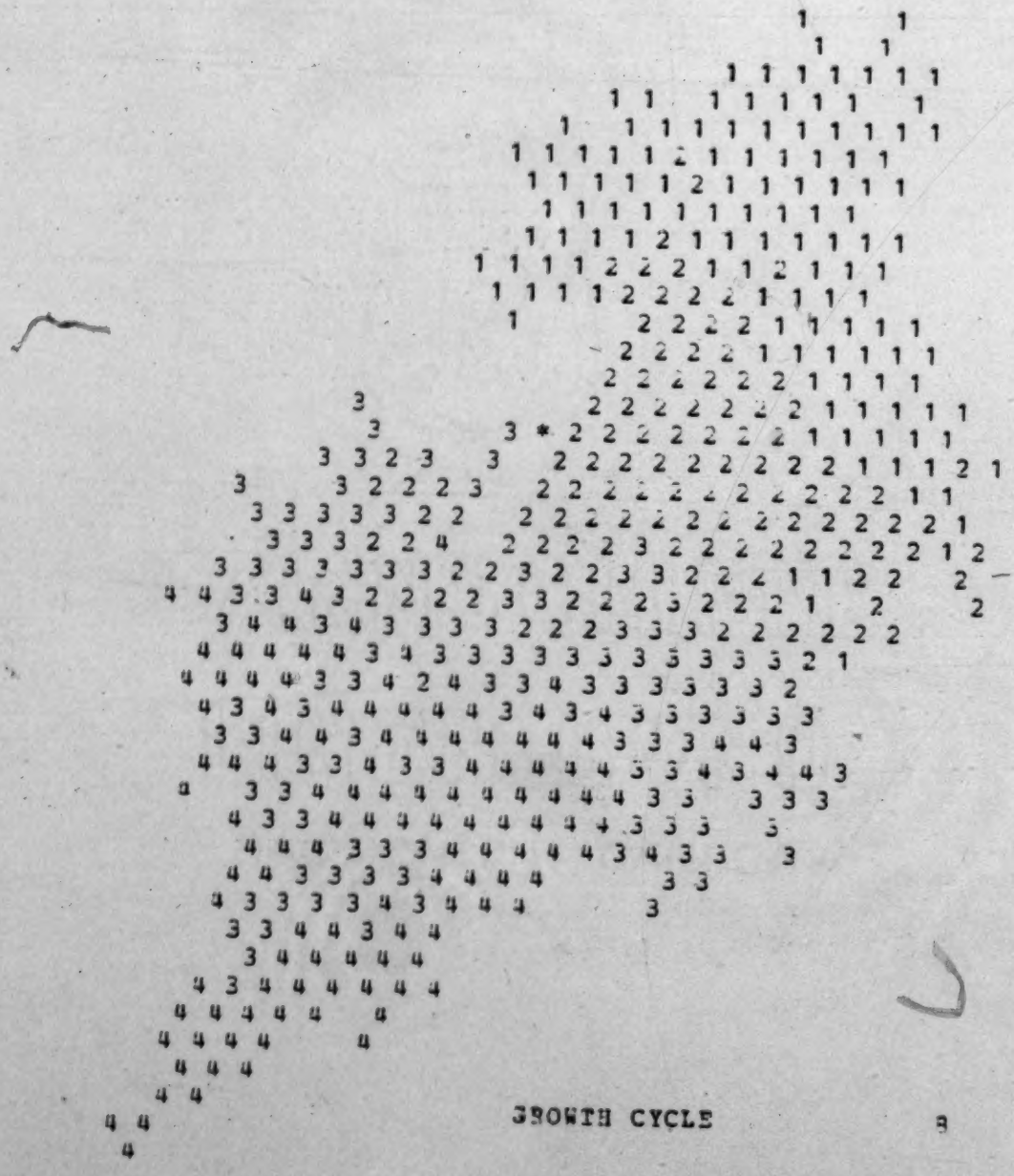


FIGURE 6.6 Growth along and around a barrier

means "uniform random" unless otherwise specified.

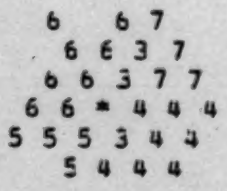
Option 1: Cell division is a determinate event occurring at regular intervals for each cell. In the model, g1 outputs a 1 whenever an internal clock of the cell (one of its state variables) reaches a maximum value (measuring the end of the cell cycle).

Output from a particular experiment using this option is presented in Figure 6.7. An initially "random" setting of clock values (positions in the cell cycle) was imposed on the cells. Cells attaining a clock value of 7, in this example, will divide in the next growth cycle, and their clocks will be reset to 1.

Option 2: Cells divide in response to the attainment of a threshold value of a local function (one of the state variables). An experiment using this option, in which the part of the function is played by f=x, where x is one of the position coordinates, with threshold value of 2, is presented in Figure 6.8 (x is the horizontal coordinate). It should be noted that this "function", which implies a cell's knowledge of its coordinates (non-local knowledge), is not intended to be realistic, but is merely convenient for purposes of demonstration.

Option 3: The control of cell division has deterministic and probabilistic elements. Once the events leading up to cell

University of Alberta



CLOCK

LINEAGE MARKER

Cells with clock value 7 will divide (doubling time 7)

CLOCK

LINEAGE MARKER

Clocks increment, cells divide

FIGURE 6.7 Successive time steps in clocked growth

```
5 8 7 6 * 2 3 4 5
```

```
                3
              3 3 4 4
              3 3 4 5 4
9 8 7 6 * 2 3 4 5 5 5 5
              3 4 4 5
                4   5 5
```

```
                3     3     3 3
              3 4 3 4 3 3 4
              3 3 4 4 4 4 4 4
                3 3 4 4 5 5 5 4 4
              3 3 4 4 5 5 5 5 4 4
9 8 7 6 * 2 3 4 5 5 5 5 5 5 5
            3 3 4 4 5 5 4 5 5
              3 4 4 5 4 4 5 5 5 5
                4 4 4 4 5 5 5 5
                  4 5 5 5 5
                  4 5 4       5
```

FIGURE 6.8 Division in response to the value of a local function

division (protein synthesis and DNA replication) are initiated, division will occur after a determined interval. The initiation of these events is apparently random, in proportion to a rate influenced by the environment, and aborted in cells which are in the "refractory period" following recent cell division. This model of growth in cell populations is discussed in Alberghina (1978) and is similar to many others (Kim, 1975; Kiefer, 1968; Wheldon, 1973).

Those intracellular controls on cell division rate which are affected by the environment are modelled here, in an intercellular context, by a single parameter, the doubling time (the time in which the population doubles in number). The doubling time is given as an integral number of time steps in this context. In ideal conditions the doubling time will be low, whereas in certain conditions (for example, lack of nutrient species) it will approach infinity (a "resting state").

The output function g1 of S1 operates here as follows:

1. Determine what proportion of cells divide in this time step (t): The number of dividing cells (nd) is (n*t/dt), where n is the number of cells and dt the doubling time.
2. Choose nd cells at "random" from the processing list. Compare their internal clocks to see if they are still in a refractory state following prior division; if so, replace them (at random).
3. Let g1 generate a 1 for these cells.

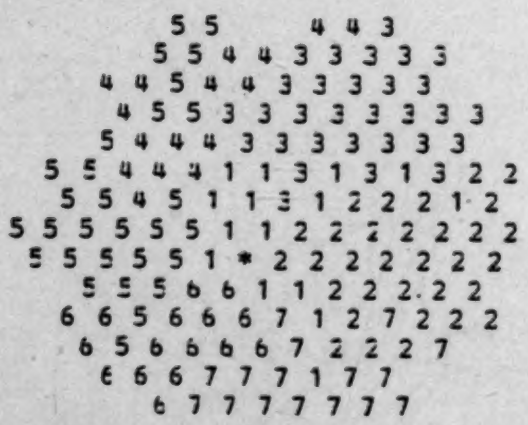
University of Alberta

Figure 6.9 illustrates growth using the random selection, doubling time, refractory period model. It should be noted that, for a doubling time of 2 time steps, the number of cells generated by the algorithm is within 5% of the expected number (n*2**k, where k is the number of times the doubling period has been repeated). The refractory period is two time steps in this example. In terms of an input-output observation morphism, the model is a valid lumped model of the Alberghina intracellular model.

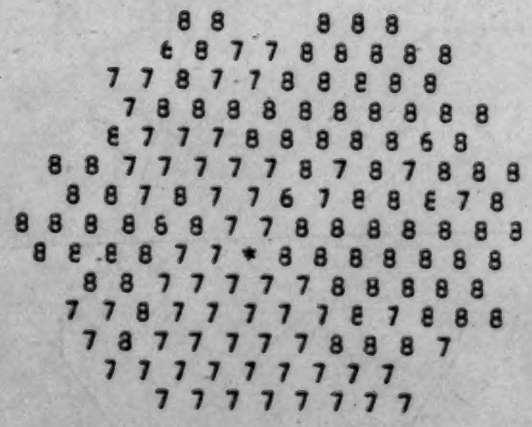
The three options of clocked growth, division in response to a local function and the Alberghina model are available and illustrate the division of individual cells in response to local conditions. Control mechanisms which have been modelled include local ones (internal cell cycle clock) and environmental input (doubling time). However, some of the central issues in controlling growth are beyond the scope of such controls. The control of shape (in formation of a limb), for example, and the limitation of cell division once a critical size has been reached, are phenomena which appear to have their domain in groups of cells and are therefore not simply explicable in local terms. Under the assumption that such phenomena appear to the individual cell in terms of local controls, the "global" control of shape and size is most probably effected by configurations arising in groups of cells due to intercellular processes or "communication".

The apparently random initiation of cell division by

University of Alberta



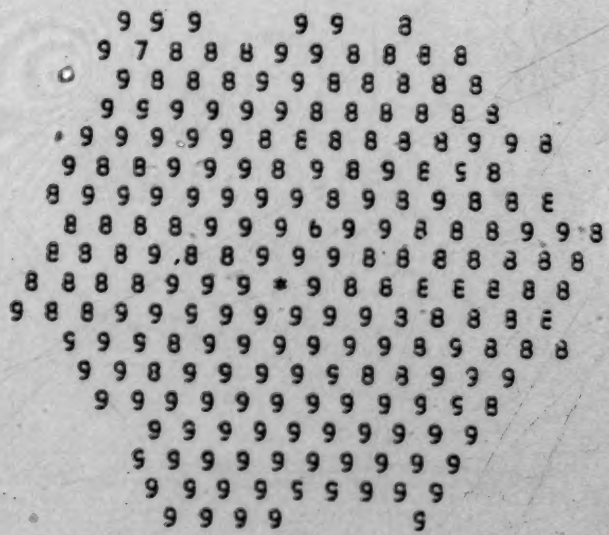
Lineage marker



Clock, doubling time 2

FIGURE 6.9a Some cells with odd clock settings will divide

Lineage marker



Clock

FIGURE 6.9b 10 growth cycles from 7 initial cells

cells not in the refractory period may hide a varying signal from an initiator substance, whose recognition is complicated by the internal clock states of the target cells. Such oscillation phenomena have been noted in organisms (Goodwin, 1973, Reiner, 1968). At the critical size such a signal might no longer exceed a threshold initiating cell division. Reaction-diffusion systems exist which generate travelling waves in response to an oscillating source; for particular values of the system parameters, such waves are transmitted unattenuated (Gmitro and Scriven, 1965). The "critical size" might be one in which unattenuated transmission is no longer supported.

In some instances capacity for cell division appears to vanish as cells become differentiated; however, the healing of wounds occurring in structures of such cells (Bryant, 1971) might indicate an ability to respond to disruption of a stable configuration inhibiting cell division. The differentiated cells either sense such disruption (perhaps the critical size is no longer operative) with renewed activity or are somehow re-enabled to divide. The latter alternative, implying an undoing of the differentiation process, appears unlikely.

Such global mechanisms of controlling growth should be susceptible to modelling in terms of division in response to a local threshold (Option 2 above), where the concentration of initiator substance is modelled in terms of intercellular communication of the reaction-diffusion type.

University of Alberta

6.2.5 Modelling the Division Direction

The shape of a growing group of cells clearly can be affected by the division direction of cells. The options mentioned in the processing experiment were a random choice of direction, division towards the direction of least pressure, and division down a concentration gradient. The problem is that of the orientation of a linearly distributed mass in the presence or absence of body forces. A lone cell might be expected to divide in a random direction in the absence of other influences; a cell in a mass of cells might be expected to experience pressure and divide in the direction of least pressure, unless more dramatic influences were present. However, if there is pronounced flow of a chemical species in a particular direction, the division axis might be expected to lie along this direction in the manner of a log floating in a current. Such an event may occur when a body of cells is entering or leaving a critical size for some reaction-diffusion system; in the transition from one stable state to another, flow of the relevant species will occur, and it is unlikely that any other systems have the same critical sizes and are in transition simultaneously.

This mechanism suggests a purely mechanical coupling between intercellular communication and the control of shape in growth. In a structure such as a limb, at non-critical phases general enlargement might be expected, while at critical values of limb length, when flow up or down the

University of Alberta

limb occurs, an elongation phase might occur. Such phases have been observed in the chick wing and other structures (Wolpert, 1978; Wilby and Ede, 1976).

6.2.6 Summary of Experimentation with Growth

The processing experiment has established that there is no apparent distortion in distribution of cell lineages (clones) due to processing order; to this extent the sequential lumped model for the parallel base model of growth is valid. Further validation will depend on further input-output observation in the context of intercellular processes. Some guidelines for time step management have been developed. Options for growth in contiguous structures have been described, as have the options of clocked cell division, division in response to the threshold of a function, and the Alberghina division decision model. Possible mechanisms for the nonlocal control of shape and size have been discussed, and an example given of shape mediated by direction of cell division under the influence of intercellular communication.

It should be noted that these simulation results cannot be compared directly with others, since they are unique with respect to the combination of individual cellular response (intracellular modelling) and changes in structures of cells (intercellular modelling). The cell space models of Wilby and Ede (1976) and Ransom (1977) described in Section 3.5, provide for individual decisions to divide, but influences

University of Alberta

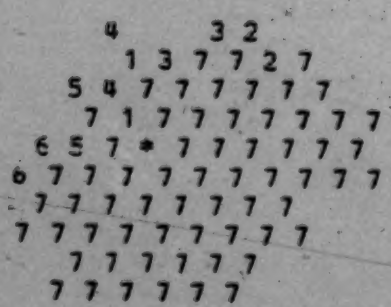
on such divisions cannot be locally mediated. Models such as that of Alberghina (1977), described in Section 3.5 and above, describe intracellular (local) influences but are not applied to a structure of cells; their domain is statistical measures on a cell population. The modelling options presented here are unique in providing the opportunity to combine both levels of modelling.

An interesting sideline of the development of the growth model is that the growth of cancerous cells can be studied. A cancerous cell is one in which division is no longer properly controlled (Kim, 1975); in the Alberghina model, this means having a doubling time smaller than the "background" cells. Figure 6.10 illustrates the accelerated division of cell lineage "7" in contrast to normal cells. Cell lineage 7 has a doubling time of 1, while background cells have a doubling time of 8 time steps. A cancer researcher might experiment with sizes of tumours as related to the relative doubling times of normal and cancerous cells, and relate his results to an intracellular model of division control.

University of Alberta

6.3 Modelling Intercellular Communication

The accumulating evidence for positional specificity of cells in developmental systems necessitates a search for processes in such systems capable of developing local values for a cell which define it relative to other cells. Chemical species involved in the competitive activities of



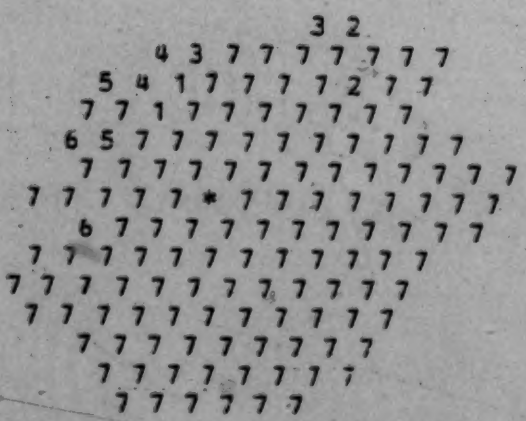


FIGURE 6.10 Lineage 7 divides without control

reaction and diffusion have been considered in Section 3.4 as mechanisms capable of producing positional specificity (pattern) from an initially uniform stable configuration.

A naive attempt to model the transmission of signals between cells was initially made, in which a proportion of the current concentration of a cell was shared between its neighbours at each time step. Whether each cell possessed in addition a reactive capacity to produce the species or not, the process inevitably produced exponentially increasing concentrations. The difference equations representing this process are unstable except for prohibitively small time steps. They may be regarded as approximations to the diffusion equation

$$u'(r,t)=L(u(r,t)),$$

where u' is the partial derivative with respect to time and $L(u)$ is the Laplacian operator with respect to the spatial coordinates r. The stability criterion for explicit approximations of such an equation is that the time step be proportional to the square of the spatial step, which makes explicit approximation impractical. Instability for larger time steps yields meaningless results as numerical errors increase exponentially with time.

Implicit methods for solving partial differential equations involve adopting a difference approximation to the differential equation in which the unknown quantity, $u(r,t+1)$, appears explicitly on the right of the equation. For each point in a spatial difference grid such an equation

University of Alberta

is specified, and the resulting system of simultaneous equations is solved at each time step. The stability of such a method depends on the eigenvalues of the matrix of coefficients of u(r,t+1); if no eigenvalue exceeds 1 the process is stable (Fox, 1962).

Reaction-diffusion systems of interest are likely to be nonlinear (Grierer and Meinhardt, 1972) and coupled (Gmitro and Scriven, 1966, Kauffman, 1978). Nonlinearity may be treated in finite difference approximations by iterative methods of solution of the resulting nonlinear difference equations; however, the analysis of departures from stable equilibria make the use of quasilinearized forms of the equations appropriate (Gmitro and Scriven, 1966).

Coupling of equations may be dealt with either by representing the coupling terms explicitly in the linear difference equations or by the use of predictor-corrector methods. The former approach necessitates m^2n^2 storage, where m is the number of coupled species and n the number of points, and mnt time units, where t is the average time to solve for one unknown if an elimination method of solution is used. The predictor-corrector approach uses much less space, mn^2, but the time for solution is Kmnt where K is the number of uses of predictor and corrector until convergence is obtained at each time step. Since the stability criterion is such that iterative methods of solution of the system of equations are effective, time is less at a premium than space and the predictor-corrector approach is selected.

University of Alberta

The Crank-Nicholson approximation (Fox, 1962) was judged appropriate for the solution of the linearized system A:

$$u1'(r,t)=K(1,1)u1+K(1,2)u2+\ldots+K(1,n)uN+D(1)L(u1)$$

$$u2'(r,t)=K(2,1)u1+K(2,2)u2+\ldots+K(2,n)uN+d(2)L(u2)$$

$$\vdots$$

$$uN'(r,t)=K(N,1)u1+\ldots+K(n,n)uN+d(N)L(uN)$$

In the approximation, the quantities on the right hand side are replaced by the average of their values at t+1 and t, the Laplacian is replaced by central difference formulae for the second derivatives, and the left hand side is replaced by a single forward difference approximation. The method has local truncation error $O(k^2+h^2)$ in regions free of singularities and therefore combines accuracy and stability.

Boundary values are incorporated into the linear system of equations, care being taken that the eigenvalues of the resulting matrix not exceed zero. Boundary values in problems of interest are either normal flux equal to zero at the boundary (no communication with outside regions) or specified values of concentrations on the boundary. Together with initial values for u(r,t) at the initial time, they ensure a unique analytic solution to the differential system, and therefore, under conditions of convergence and stability, to the difference system.

University of Alberta

6.3.1 Formal Specification of the Communication Model

A communication model is a structure

S=<Q,T,I,W,d>, where

Q is the single state variable, consisting of a linked list of cells specified by neighbourhood links and pointers to individual parameters;

T is the continuous time base;

I is the input value set, incorporating communication through the boundaries of the structure with the environment (boundary values);

W is the input segment set, incorporating the possible dependence of the boundary values on time;

d is the state transition function, d:Q x I --> Q

The state transition function d generates the new state q by changing one or more of the parameters pointed to by each cell. Its domain is the entire structure, since change in each cell depends on the state of neighbouring cells. In the present instance, d is given in differential form in a reaction-diffusion system for the concentration of one or more chemical species.

The continuous time system S can be simulated exactly by a discrete time system SD provided that its state trajectories are to be reproduced only at sample intervals of constant length h (Ziegler, 1976). However the trajectories can be generated only if the global transition function (that is, the analytic solution of the differential system) is known. SD is therefore unobtainable as a

University of Alberta

practical basis for iterative generation of state trajectories, but may be used as a standard for judging approximate methods.

The system SD is given by

SD=<Q,TD,I,Wh,dh>, where

Q, I are as in S;

TD is a discrete time base;

Wh is the input segment set W restricted to segments of length h;

dh is the state transition function operating at the endpoints of input segments (at discrete time steps nh).

Corresponding to the system SD there is an iterative specification (Ziegler, 1976), and corresponding to the iterative specification a sequential machine Md which simulates it exactly (generates the state trajectories given the global transition function). Let MD(h) be given by <Wh,Q,dh>. Now an approximation method, which represents the differential system as a system of difference equations, may be referred to as

MA(h)=<I,Q,da(h)>, where the transition function

da(h):Q x I --> Q

is given by the solution of the difference system for q(t+h).

It is necessary to establish an "approximate" specification morphism between MA(h) and MD(h) in order that the lumped numerical model MA(h) may be regarded as a valid simulator of the lumped model MD(h), and therefore of the

University of Alberta

base model S in a particular experimental frame. A specification morphism between the two sequential machines (preservation of transition function) is established by proving that the trajectories generated by the approximate machine MA(h) come arbitrarily close to those generated by MD(h) on identical input segments.

Conditions under which this situation, known as "convergence" in numerical mathematics, will hold true depend on "accuracy" and "stability". The accuracy of a numerical method depends on its local truncation error; that is , on

$$|dh(g,w)-da(h)(g,w(0)|=M$$

In the case of the Crank-Nicholson method, the truncation error is $O(h^2+k^2)$. The stability of a numerical method depends on the propagation of this error, and the errors incurred by the use of finite mathematics, by repeated application of the method. In the Crank-Nicolson approximation the eigenvalues of the matrix of coefficients in the difference equations are less than one, which implies that repeated solution of the system does not allow unbounded error amplification (Fox, 1962). In terms of MD(h) and MA(h) this means the accumulated error after n steps

$$|dh(g,w1,w2,...,wn)-da(h)(g,w1(0),w2(0),...,wN(0))|$$

is bounded by Mp(h) where p is a polynomial expression in h. Convergence is then assured as Mp(h) --> 0 as h --> 0, and the error is arbitrarily small.

University of Alberta

Thus, because the characteristics of the Crank-Nicholson approximation are known, it may be asserted that an approximate specification morphism holds between MA(h) and MD(h) (between the numerical, iterative solutions of the difference equations and the analytic solutions of the differential equations) in any experimental frame.

The real significance of this fact is that experiments run by the simulator generating trajectories may be considered direct tests of the base model (reaction-diffusion) in comparison with trajectories of the real system in a particular experimental frame; that is, the effects of replacing the differential system by a difference (spatially and temporally discrete) system may be ignored. Experiments with reaction-diffusion as the mechanism of communication between cells may therefore concentrate on observing generated trajectories (a sequence of "patterns") and comparing them with events in real systems, thus investigating the equivalence of their input-output relations (postulate 10, Ziegler (1976)).

It may be noted that the sequential machine MA(h) is identified directly with the simulator or computer program. MA(h) operates in parallel on the cell structure, updating all cells simultaneously for each time step. The implicit equations are solved as a simultaneous system in the computer program. Parallel processing is thus preserved and the identification is justified.

The communication model is specified only at the level

University of Alberta

of the cell structure, in contrast to the growth model.

6.3.2 Experiments with Numerical Solutions of Reaction-Diffusion Systems

The concept of positional specificity in prepatterning cells for differentiation is a hypothesis inferred from many events in development, as discussed in Chapter Three. It has proved impractical to obtain direct evidence, in the form of local readings on concentration values, of such fields due to the multitude of processes occurring in cells and the difficulty of performing in vivo experiments. The "real system", therefore, cannot be given by concentration data from a functioning morphogenetic field. For the purposes of demonstrating the efficacy of reaction-diffusion as a mechanism for generating positional specificity, the real system will be considered to be any group of cells displaying such specificity, that is, nonuniformity.

Prepatterning hypothesis: Reaction-diffusion is a mechanism capable of producing stable nonuniform configurations from uniform configurations.

In order to provide a background for experimentation with the prepatterning hypothesis a brief review of the

University of Alberta

instability analysis for R-D systems is justified. Reaction and diffusion are "competing" processes in that reaction is disruptive (produces local change) whereas diffusion is dissipative (distributes irregularities). This interaction may be expressed in nonlinear (coupled) differential equations by means of the relative magnitude of the reaction rates and diffusion rates. The solution of such equations may be expressed as the sum of a uniform, steady state solution and an excursion from that state. The differential system may then be adjusted to describing the excursion from steady state by means of the process of quasilinearization about the steady state (Gmitro and Scriven, 1966). The spatial dependence of the solution for the excursion in a domain may be expressed as the sum of the Laplacian shape functions relevant to that domain, multiplied by an exponential growth factor related to the relative magnitudes of the reaction and diffusion parameters (Gmitro and Scriven, 1966, Kauffman et al, 1978). In general this sum is zero and the steady state is maintained; however, at critical sizes of the domain, one particular shape function may fit exactly into the domain; it is then selected for amplification and causes a departure from uniform steady state.

Gmitro and Scriven (1966) discuss qualitatively criteria under which such events occur in systems of one, two or three participating species. For N participating species such behavior can be ensured by only (N*N+3*N)/2

University of Alberta

constraints on 2N*N+1 parameters. Thus there is a considerable degree of freedom in selecting values for the reaction and diffusion parameters. In terms of the real system this implies versatility, but in terms of experimentation it implies the impracticality of exhaustive investigation of behaviors of reaction-diffusion systems. However, for the purposes of testing the prepatterning hypothesis, it is sufficient to establish such behavior in only one system. The system chosen for investigation is that of Kauffman et al. (1978) given by

u1' =K(1,1)u1+K(1,2)u2+D(1)L(u1)

u2' =K(2,1)u1+K(2,2)u2+D(2)L(u2)

where K(1,1)=-7.8, K(1,2)=13.4, K(2,1)=-1, K(2,2)=2.04 D(1)=16, D(2)=1.

Analysis by the methods of Gmitro and Scriven (1966) predicts that the critical size range is the interval (6.1,9.1) (Kauffman et al. 1978).

Prepatterning experiment: To establish the prepatterning hypothesis by observing departures from uniform steady state in a reaction diffusion system, namely the Kauffman system given above.

Experimental frame: The Kauffman system is to be applied in one- and two-dimensional rectangular domains. The output trajectories for various sizes are to be compared to the expected shape functions. The boundary conditions are that

University of Alberta

the normal flux be zero, corresponding to an autonomous system (no communication through the boundary); the shape functions expected for rectangular domains under these boundary conditions are integral numbers of half-cosine-waves (Kauffman et al., 1978).

Procedure: The Crank-Nicolson method is used in generating trajectories in one- and two-dimensional domains. In the case of one dimension, successive critical sizes are investigated by manipulating h, here representing the spatial step size or distance between cell locations, and n, the number of cells, bearing in mind that h must be less than 1 to ensure accuracy. In the case of two dimensions, since the Crank-Nicolson method, with its stability characteristics, is represented in a rectangular coordinate system, one of two actions is necessary. Either the Crank-Nicholson method must be rephrased in terms of a hexagonal array, or the hexagonal array must be converted to a rectangular one. The second alternative is chosen, firstly, because the Laplacian operator in hexagonal coordinates exhibits mixed partial derivatives which prohibit the use of alternate-direction acceleration techniques (Fox, 1962) which may prove useful in some circumstances, and secondly, because there is no guarantee of stability in the resulting coefficient matrix. Conversion of the hexagonally arranged cells into a rectangular array for the purposes of reaction-diffusion in some of their parameters necessitates the

University of Alberta

initial interpolation of extra values for these parameters at intermediate points. Aitken interpolation in two directions is used to achieve this, care being taken that enough points are used to maintain the accuracy of the Crank-Nicholson method. When a trajectory is completed, the extra values are discarded and the concentration values returned to the parameter records of the corresponding cells.

Results: The expected patterns were observed in the expected size ranges, and are presented in Figures 6.11-6.14.

Discussion: The trajectories generated by the reaction diffusion equation display a dependence on various parameters, both of the numerical approximation and of the differential system itself.

Refinements of the spatial grid and of the time step produce faster convergence of the predictor-corrector pair used to represent the coupling of the two-species equation, but increase the order of the system to be solved and number of time steps to be taken; an optimal balance has to be struck, minimizing the number of arithmetic operations. The unconditional stability of the method is demonstrated by the freedom with which such manipulation can be performed with regard only to the accuracy constraint that the steps be less than 1.

The speed with which the patterns reach within 0.1 of

University of Alberta

SPECIES 1 PLOTTED FOR 1 OUT OF 1 CELLS
CONCENTRATION SCALED BY 1.000000
SECTION ALONG X AXIS
TIME STEP 100

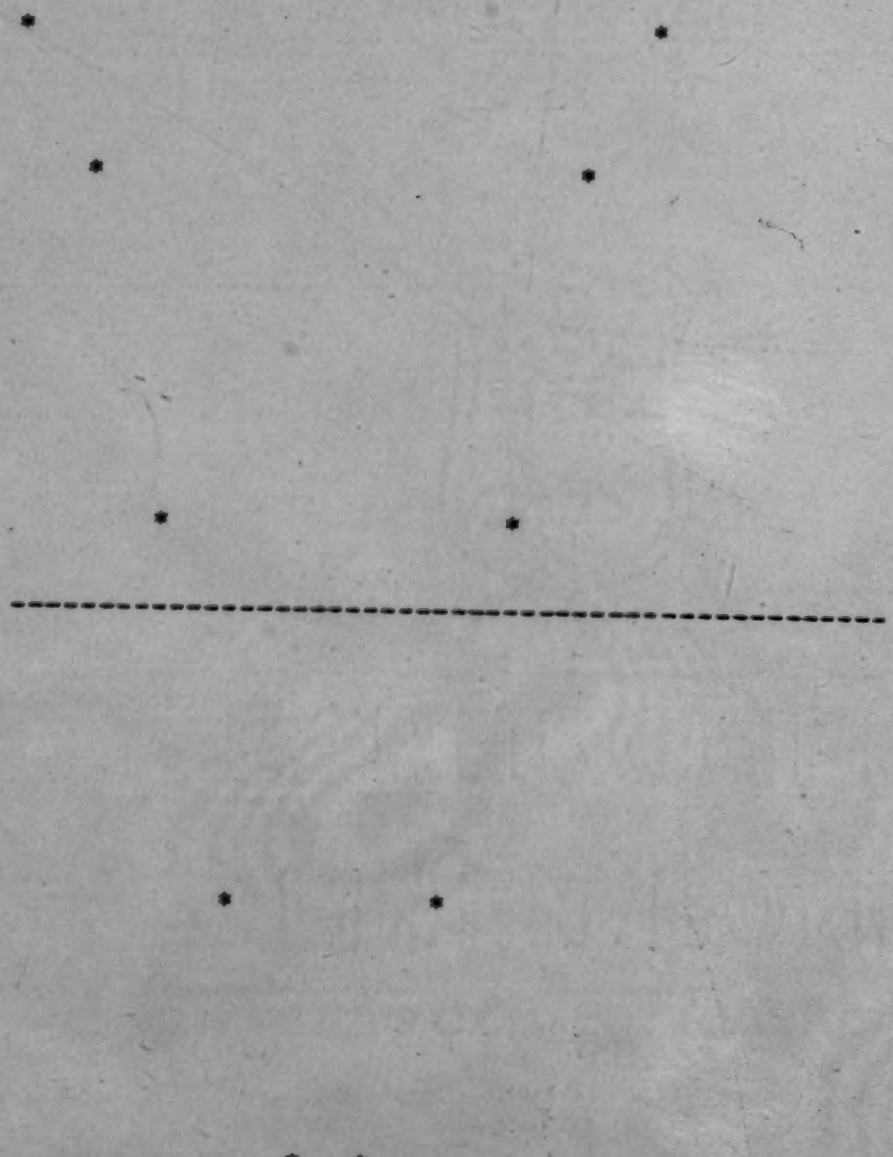
FIGURE 6.11a First cosine pattern arises in one dimensional, two species system

SPECIES 1 PLOTTED FOR 1 OUT OF 1 CELLS
CONCENTRATION SCALED BY 1.000000
SECTION ALONG X AXIS

FIGURE 6.11b First cosine pattern established, species one

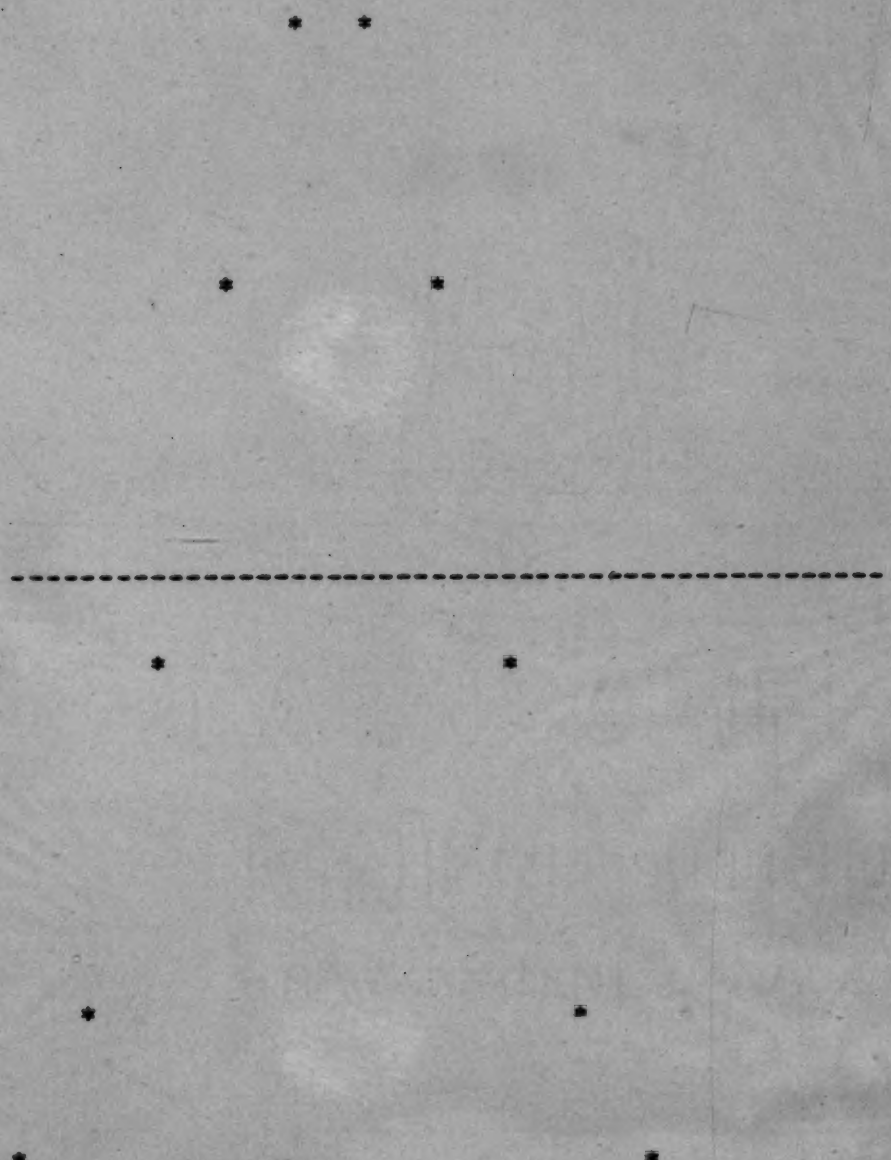
SPECIES 2 PLOTTED FOR 1 OUT OF 1 CELLS
CONCENTRATION SCALED BY 1.000000
SECTION ALONG X AXIS

FIGURE 6.11c First cosine pattern, species two



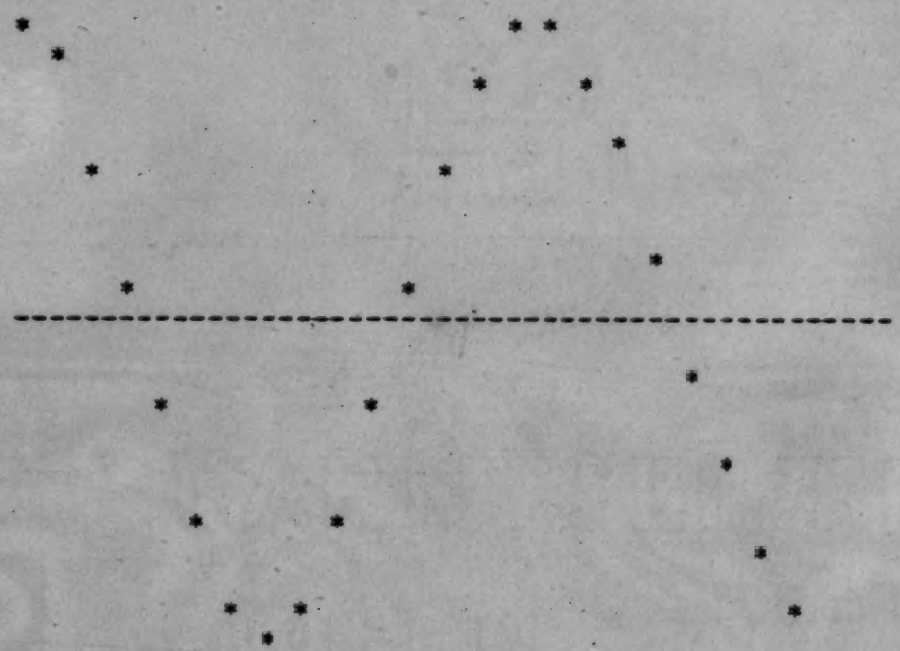
SPECIES 1 PLOTTED FOR 1 OUT OF 1 CELLS
CONCENTRATION SCALED BY 1.000000
SECTION ALONG Y AXIS

FIGURE 6.12a Second cosine pattern, species one



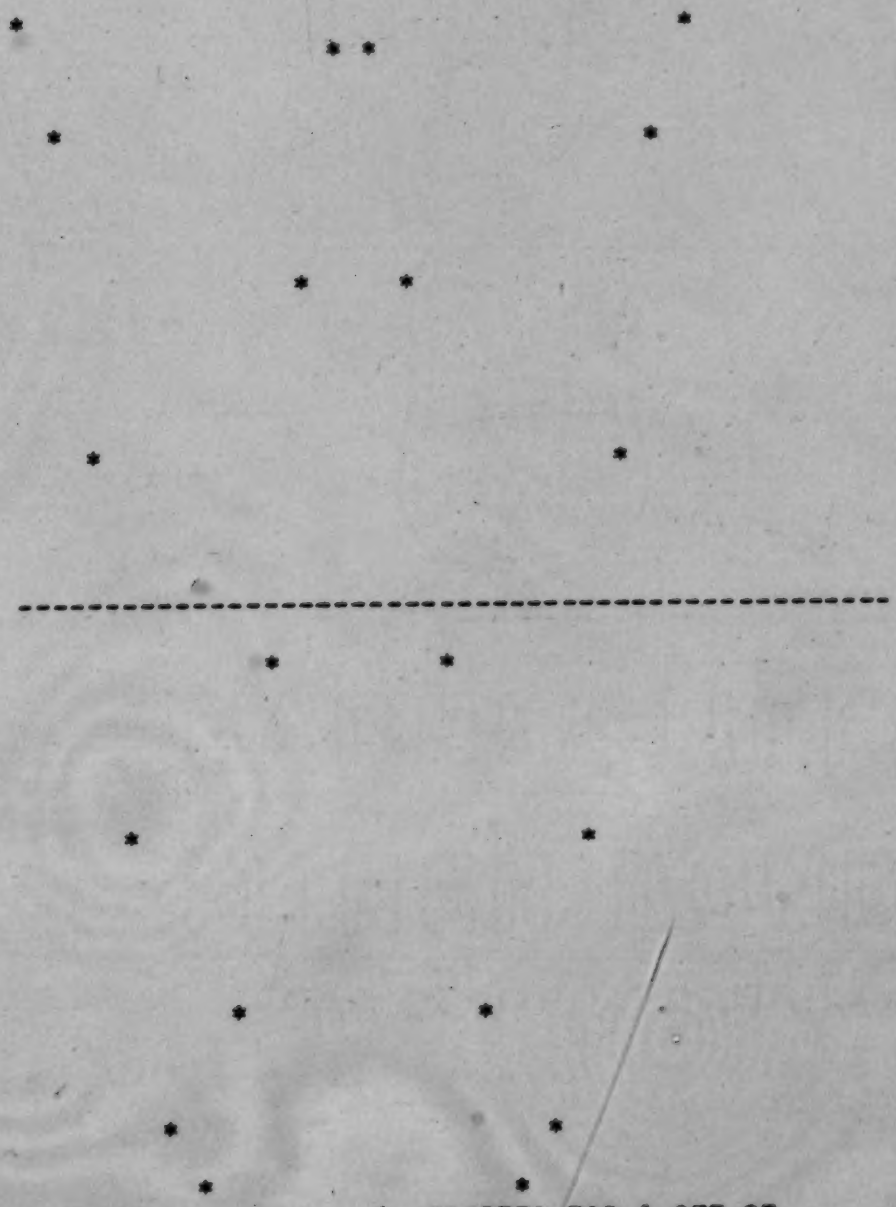
SPECIES 2 PLOTTED FOR 1 OUT OF 1 CELLS
CONCENTRATION SCALED BY 1.000000
SECTION ALONG X AXIS

FIGURE 6.12b Second cosine pattern, species two



SPECIES 1 PICTTED FOR 1 OUT OF 1 CELLS
CONCENTRATION SCALED BY 1.000000
SECTION ALONG X AXIS

FIGURE 6.13 Third cosine pattern, species one



SPECIES 1 PLOTTED FOR 1 OUT OF 1 CELLS
CONCENTRATION SCALED BY 1.000000
SECTION ALONG X AXIS

FIGURE 6.14 Fourth cosine pattern, species one

the pure cosine functions varies with the initial disturbance (noise) induced locally in the uniform steady state to provide the impetus for departure from uniformity. It may be noted that when noise is not modelled directly, the accumulating numerical error is sufficient eventually to initiate the formation of the non-uniform pattern. The system's effect on numerical error is interesting in general. In uniform (noncritical) states, the Laplacian term "distributes" the error in the system, thus dissipating it; at critical sizes, numerical error acts constructively in establishing the pattern, and then in the new uniform configuration is again dissipated. A speculation may be advanced as to whether the discrete representation of reaction-diffusion perhaps has a structural validity for the real system (that is, that cells should be regarded as discrete entities and not as points in a continuous differential field).

A disturbance in only one cell of a field is sufficient to induce the pattern. Gmitro and Scriven (1966) suggest that noise in the real system is caused by "thermal fluctuations" in the concentrations due to constituent molecule kinetics. In a developing system, however, another source of variation is provided by dividing cells. During mitosis, copies are made of some substances while others are distributed, perhaps not evenly, between daughter cells. A local disturbance in concentration results. It may well be that cell division, taking a morphogenetic field from one

University of Alberta

pattern to the next as the size increases, also controls the speed with which patterns arise and decay by providing noise of a certain magnitude. If so, it is another instance of the essential interdependence of growth and determination in developmental systems.

The system explored in the prepatterning experiment may be thought to be highly artificial; in reality there is such a high degree of freedom in the system that it may be used to represent many actual configurations. Firstly, critical lengths may be achieved in fields of any desired size by varying the step size (distance between cells) in inverse proportion to the number of cells. Secondly, the K and D parameters are chosen to satisfy certain conditions for the desired stability behavior (Gmitro and Scriven, 1966), such conditions in this case imposing 5 constraints on 6 quantities. The remaining degree of freedom permits the system parameters to be scaled downwards to values realistic for cell reactions and diffusions. The experimenter may take one of two approaches; he may postulate values for the parameters in the study of a particular real morphogenic field, and see whether the resultant patterns are established after a realistic time period, or he may use knowledge about the actual time taken in establishing a field in order to "calibrate" the system parameters.

While the Kauffman system cannot, obviously, represent all stability behaviors of interest in pattern formation models, it does, because of the discussed degrees of

University of Alberta

freedom, have the capability to represent (at least qualitatively) events occurring in many developing systems. This property of the system is used to advantage in Section 6.5.

Conclusions: Experimentation has established the prepatterning hypothesis, that is, reaction-diffusion is a mechanism capable of producing nonuniform configurations from uniformity.

The use of numerical techniques becomes costly with many points, particularly in two-dimensional systems where there is translation to a rectangular grid. The question arises as to whether, in domains whose shape functions are known, iterative generation of the trajectories in the growing domain is necessary. Under certain conditions it is. If the appearance and decay of the patterns are to be studied for themselves, or if other processes being modelled depend dynamically on local concentrations, then iterative generation is essential.

University of Alberta

6.3.3 Modelling Using Analytical Solutions

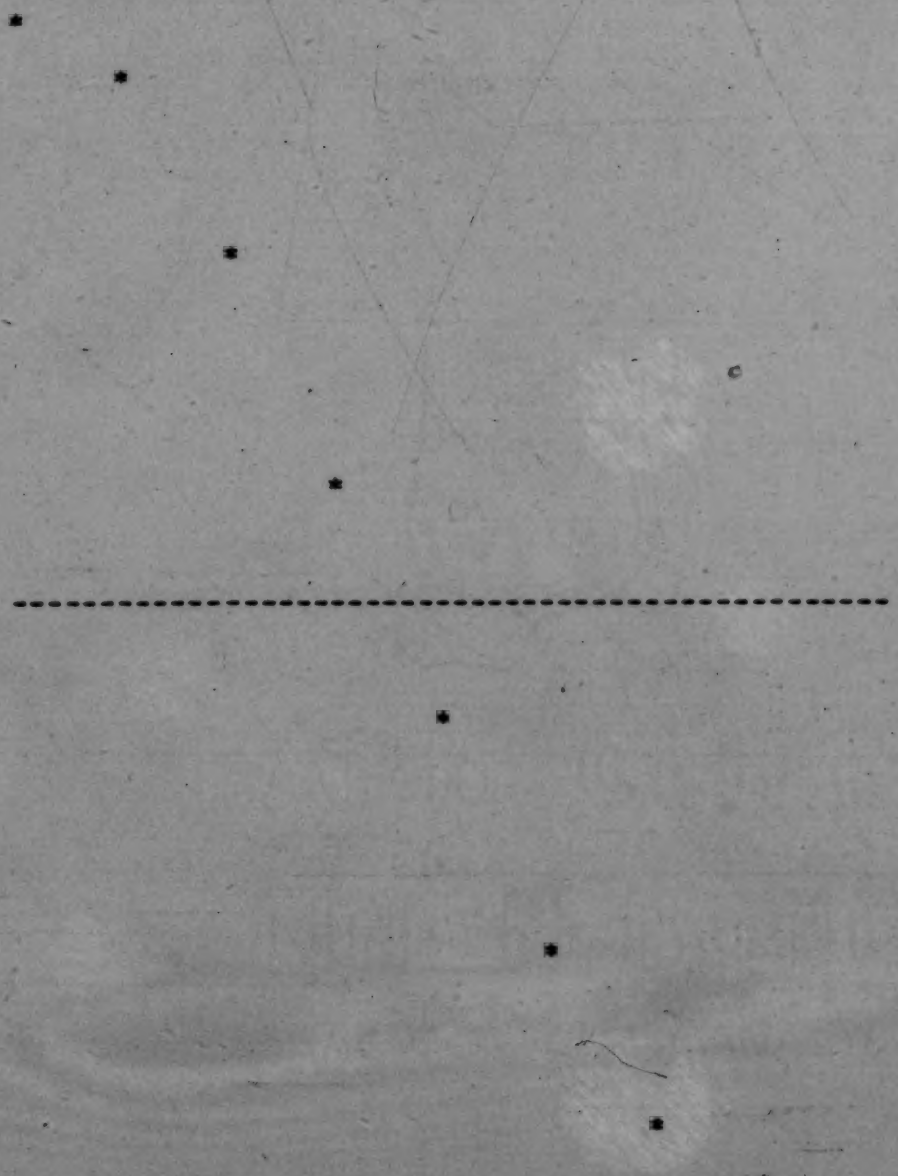
In some modelling situations the communication model may be independent of any other model and the shape functions for the system may be available. The reaction-diffusion continuous system S may then be modelled by the machine MD(h) described in Section 6.3.1, the exact solution machine giving trajectories at discrete values of time. The

time base of MD(h) has variable step length, the length of any one step being given in the continuous system by the time between establishment of one configuration and the next. An "event" in MD(h) is initiated when critical size is reached, that is, when a dynamically mediated condition on the structure of the cell system is satisfied.

The option of using analytic solutions for reaction-diffusion systems in one-dimensional and two-dimensional rectangular domains is provided. The relevant shape functions are sines or cosines, as is demonstrated in Figures 6.15 and 6.16. In Figure 6.15a and 6.15b, the solutions for one of the species in orthogonal directions are given, and in Figure 6.15c the result of multiplying these components is shown by indicating the sign at individual cells in the two-dimensional structure. Since these patterns of signs at individual cell indicate the location of lines where the solutions are zero, they are referred to as "nodal patterns". Similarly Figures 6.16a-c show the solution for another critical size.

The shape functions for circular and elliptic domains are Bessel and Matthieu functions respectively (Kauffman et al., 1978). The nodal patterns which may arise given certain critical sizes of the radii, (or the major and minor axes), are known qualitatively (diameters and concentric circles, and axes, hyperbolae and confocal ellipses), but finite expressions for evaluating them, thus locating and timing the rise of nodal lines, are difficult to obtain.

University of Alberta



SPECIES 1 PLOTTED FOR 1 OUT OF 1 CELLS
CONCENTRATION SCALED BY 1.000000
SECTION ALONG X AXIS

FIGURE 6.15a Analytic solution, x direction (pattern one)

SPECIES 1 PLOTTED FOR 1 OUT OF 1 CELLS
CONCENTRATION SCALED BY 1.000000
SECTION PERPENDICULAR TO X AXIS

FIGURE 6.15b Analytic solution, z direction (noncritical)

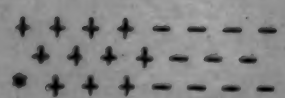


FIGURE 6.15c Nodal pattern

SPECIES 1 PLOTTED FOR 1 OUT OF 1 CELLS
CONCENTRATION SCALED BY 1.000000
SECTION ALONG X AXIS

FIGURE 6.16a Analytic solution, x direction (pattern one)

SPECIES 1 PLOTTED FOR 1 OUT OF 1 CELLS
CONCENTRATION SCALED BY 1.000000
SECTION PERPENDICULAR TO X AXIS

FIGURE 6.16b Analytic solution, z direction (pattern one)

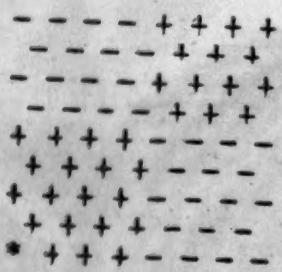


Figure 6.16c Nodal pattern

Sometimes a nodal configuration may be inferred by extrapolation from rectangular configurations or by symmetry considerations, as will be seen in Section 6.5. In a model where the positional specificity of cells is specified completely by these nodal lines, analytic expressions for the shape functions themselves are not necessary, merely knowledge of their sign.

6.3.4 Summary of Experiments with Communication

The numerical solution of reaction-diffusion systems has been investigated for a system of wide applicability. The method generates trajectories approaching those of the base model in all chosen experimental frames, reflecting the approximate specification morphism between the discrete approximation and the analytical solution. The validity of the reaction-diffusion model for producing prepatterns in morphgenetic fields must be tested in the context of observable developmental events (for example, cell differentiation) in models which address the problem of cell determination in response to pre-patterns.

The options of the use of analytic solutions or node location are provided for certain domains and situations.

6.4 Modelling Growth, Communication and Determination

Development is a process in which growth and communication between cells operate simultaneously, and critical conditions (presumably local) at various stages of

University of Alberta

the process initiate changes in the organism resulting in differentiated cell types. In order to model development, the models of growth and communication must be merged, and a model developed for recognizing critical conditions.

The base models for growth and communication, G and S, have continuous time bases and the same state set, the set of cell structures. It is therefore a simple matter to merge them into one system. However for purposes of simulation their discrete lumped models in particular experimental frames must be considered. In the same experimental frame, the time bases for the two lumped models must be constructed so as to ensure that possible influences of one process on the other are taken into account.

In Section 6.2 the possible influence of local conditions on a cell's capacity to divide was discussed. The discrete representation of the growth time base is valid if local conditions affecting growth do not change during a growth time step. That is to say, if G depends on S then timestep(G) is less than or equal to timestep(S), and timestep(S) must be an integral multiple of timestep(G).

The effect that growth has on the communication process may be negligible or profound. If one or two extra cells in a structure do not take the system into any new critical size range, the spatial patterns will not change; however, if the expansion does, a completely new configuration will arise as concentration levels adjust by means of chemical flow through the system. Since the communication process

University of Alberta

must be sensitive to the possibility of entering a new critical size at all times, S depends on G. Timestep(S) should be less than or equal to timestep(G) and timestep (G) must be an integral multiple of timestep (S).

Therefore if both processes are dependent on each other, they must have the same discrete time step. In cases where growth is independent of the communication process, communication time steps can be smaller than growth time steps.

Determination, the genetic commitment of a cell to a future path leading to differentiation, is a response to attainment of critical values in a cell's local conditions. Change in such conditions is generated by the communication model. A model for determination, therefore, being sensitive to any changes as they arise, should have the same time step as the communication model.

University of Alberta

6.4.1 Formal Specification of the Determination Model

A determination model is a structure

D=<Q,T,e> , where

Q is the single element state set, consisting of a linked list of cells specified by neighbour links and pointers to individual parameters;

T is the continuous time base;

e is the state transition function,

e:Q --> Q.

The system D has state transition function e which (a)

leaves the structure unchanged if no (local) critical parameters are found, and (b) changes the character of cells in the structure for which critical parameters are found (by manipulating the cell's parameter record, for example giving it a new "genetic activation code").

It may be remarked that determination, once initiated, is a purely local event, occurring in individual cells as an isolated response, neither changing the structural arrangement of the group of cells nor changing intercellular communication. Providing the time base is such that local conditions do not change during the process of determination, it is immaterial whether the cells are treated sequentially or in parallel by the transition function e. D may therefore be modelled by

DS = <Q,TD,ed>, where

TD is a discrete time base;

ed is the state transition function;

ed:Q --> Q, where ed is given by the following algorithm:

1. Set pointer P to the first cell on the processing list.
2. Inspect relevant parameters of the cell for critical values.
3. If a critical value is found, effect the resulting changes in the cell's character.
4. Set P to the next cell on the processing list.
5. If P is not null, go to 2, otherwise stop.

The particular changes in a cell's character brought about by determination will depend upon the experimental

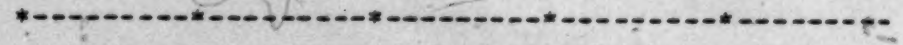
University of Alberta

frame within which a particular experiment with DS is being undertaken. It may take the form of adding digits to the developmental genetic code (see Section 4.1), or noticing a significant accumulation of such digits and initiating actual differentiation; the latter situation may result in division of the structure of cells into substructures, which possibly do not communicate similarly. The reaction-diffusion system in progress may be eliminated in the differentiated cells. Thus a still higher degree of control is implied in a model of differentiation. In Section 6.5 such systems will be discussed.

University of Alberta

6.4.2 Growth and Communication

An experiment has been conducted with one-dimensional structures in which cell division by the Alberghina model occurs simultaneously with communication using the Kauffman system. Two situations are considered: (a) the patterns arising are allowed to stabilize before further cell division is permitted (timestep (S) much less than timestep (G)), and (b) cell division is allowed to occur during the establishment (and decay) of patterns (timestep (S) less than timestep (G)). Representative results may be seen in Figure 6.17, where a one-dimensional structure is growing and undergoing various non-uniform configurations. It is found that the effect of dividing cells on the dynamics of pattern formation is as expected by Section 6.3.2. Cell division (a) helps a new pattern to rise by providing noise"

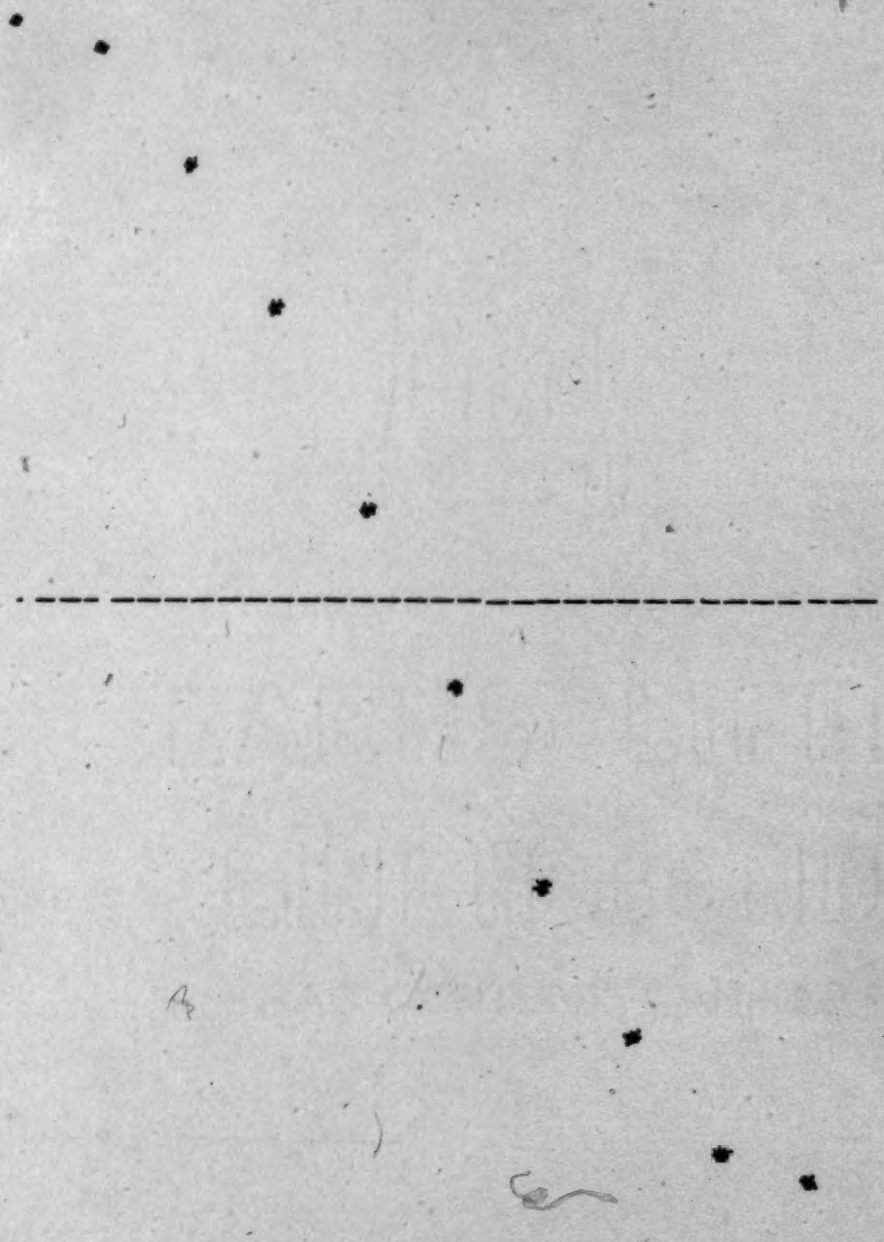


SPECIES 1 PLOTTED FOR 1 OUT OF 1 CELLS
CONCENTRATION SCALED BY 1.000000
SECTION ALONG X AXIS

•2345

lineage

FIGURE 6.17a Growth through cosine patterns, initial configuration



•122334455

lineage

FIGURE 6.17b With 10 cells

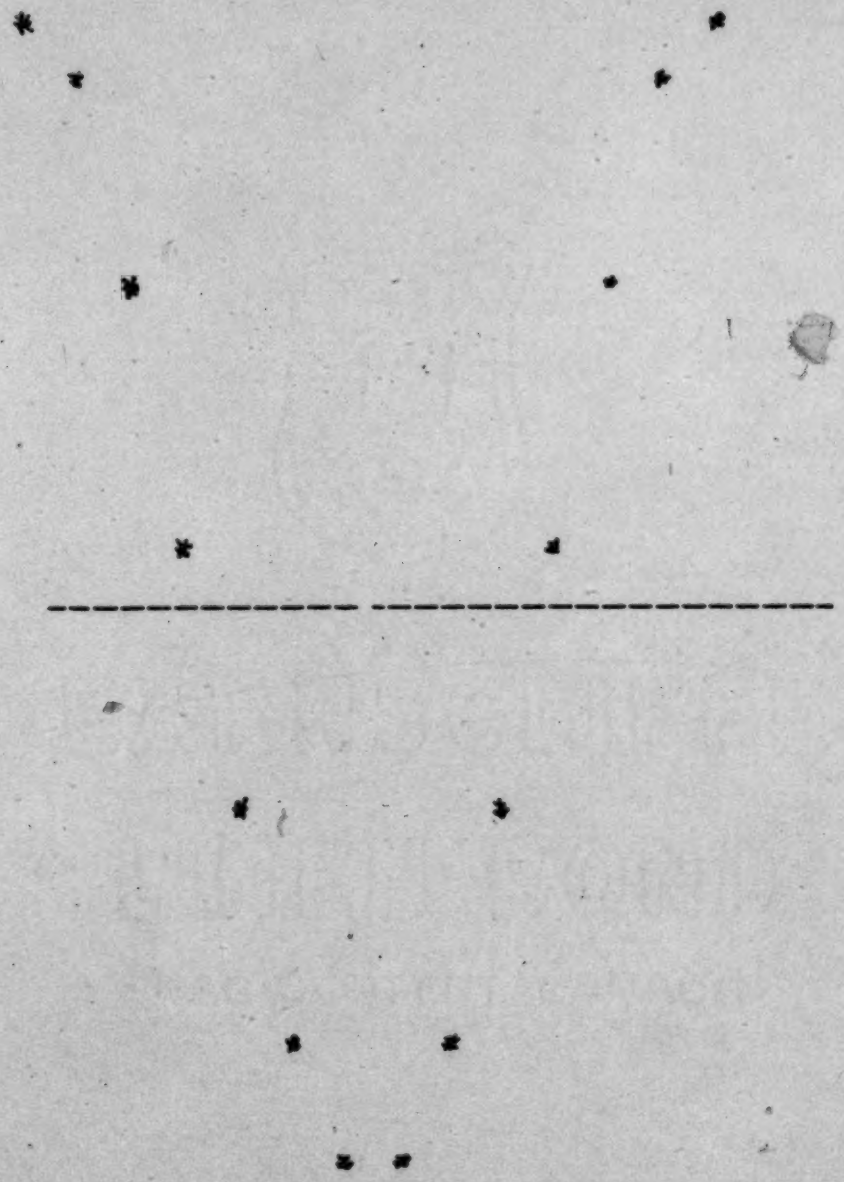
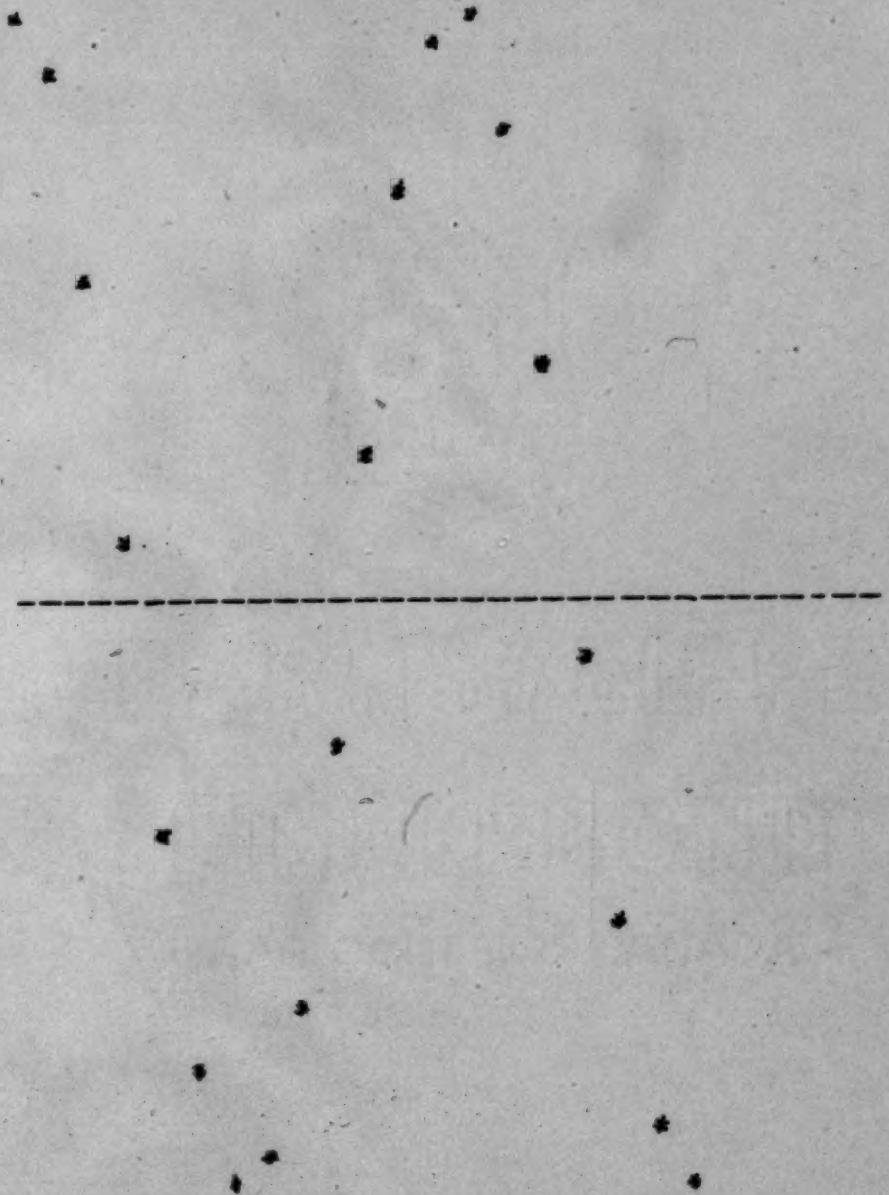


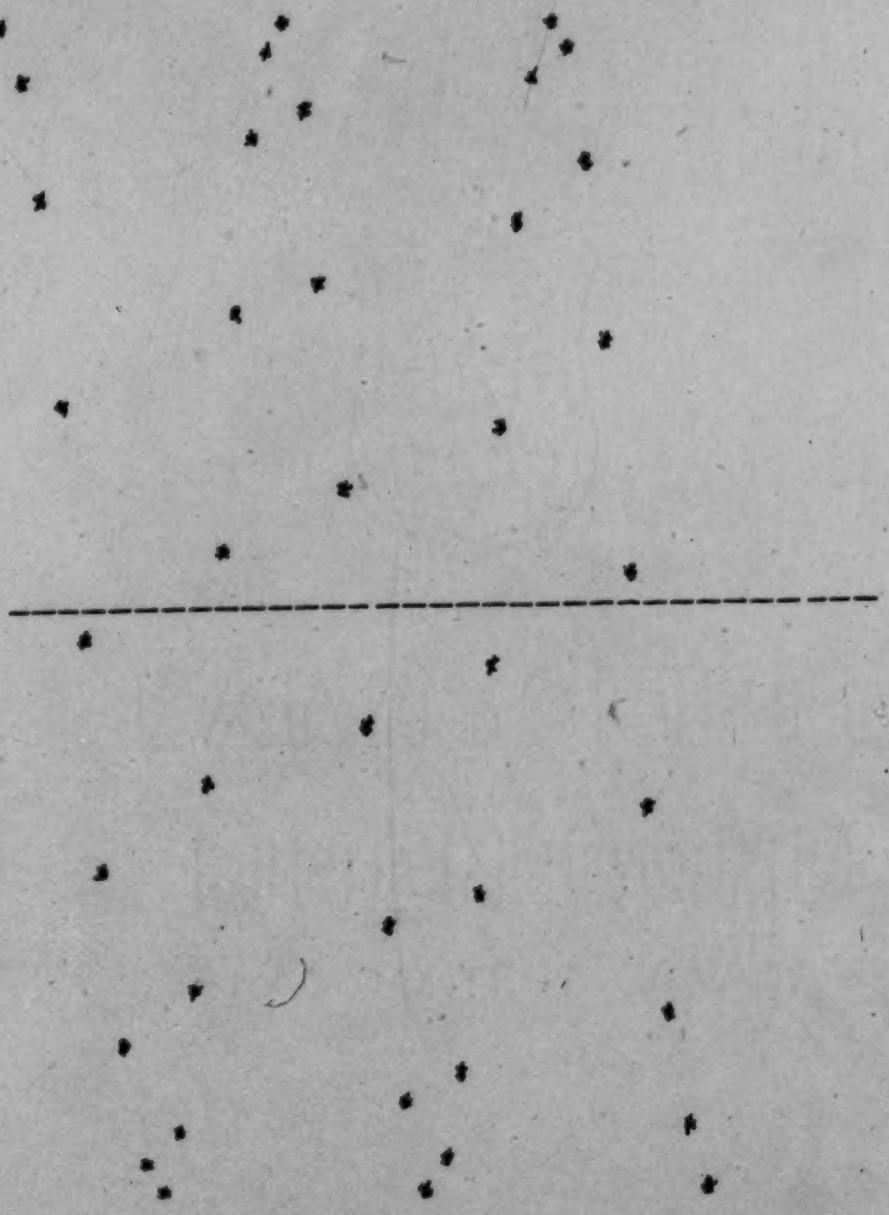
FIGURE 6.17c With 14 cells



*11112222333344445555

lineage

Figure 6.17d With 20 cells



:1111112222222223333333344444455555555555

lineage

FIGURE 6.17e With 40 cells

(simulated by sharing the concentration of the relevant substance unevenly), (b) does not affect the time taken in establishing the pattern (the local "flattening" is compensated for by the additional noise supplied to the process) unless the critical size for the system is exceeded, and (c) hastens the decay of patterns by local "flattening".

It should be remarked that the rate of growth has a direct effect on the patterns appearing in a system. If the system grows fast it may miss some of the patterns (for example, by going from 10 to 20 cells in Figure 6.17, thus missing an intermediate pattern). In such cases patterns begin to arise but the system exceeds the critical size before the pattern stabilizes. A comparison may be made between the capabilities exhibited in Figure 6.17, and the capabilities of systems like CELIA (Herman et al., 1974) which also investigates one-dimensional structures (see Section 3.4.1). Firstly, the state set for this author's model may be extensive and of mixed type, while the state set of CELIA has a single component, a discrete genetic marker. Secondly, in CELIA signal transmission is of a primitive nature. Thirdly, due to this primitive signalling, gradient patterns are achieved only by imposing special states at the end points of arrays, thus violating strict initial uniformity. In summary, CELIA presents a much less "realistic" picture of developing one-dimensional systems.

University of Alberta

6.5 Modelling Developmental Systems

A developmental system is one in which the various intercellular processes interact in a manner resulting in organized changes in cells and cell structures. In the context of this project, these processes have been classified under the headings growth, communication and determination. Development of the system appears to occur by means of coordination and control of these processes.

It would be easy to postulate a "global control" which performs the task of coordination, making sure that the growth rate takes the system through the right patterns, scheduling determination, and so on. In other words, it would be easy to specify a master system invoking growth, communication and determination when required. However to do so would be to act in contradiction of the basic assumption that development is mediated at the local level, though its effects are widespread.

How is it, then, that such processes act "cooperatively" to produce a viable system? The answer must lie in the process of natural selection: many nonviable interactions must have been discarded in selecting viable systems for survival. It is possible to imagine the genesis of a simple organism, for example. A capacity for a certain reaction-diffusion system is inherited or provoked by mutation in a cell. It divides and happens to reach a critical shape for that R-D system. The resulting nonuniform pattern triggers changes in some of the cells.

University of Alberta

This configuration is, in some sense, viable in its environment and therefore is selected for and continues to develop. That reaction-diffusion system, after surviving many such viability tests, becomes one of the developmental tools of the strain's evolutionary descendents.

In combining all of the processes of development in order to model a particular real system, then, what appears to be a controlled, cooperative interaction must be thought of, rather, in terms of fortuitous selection of system parameters. The growth and communication time bases must be selected with reference to the particular experimental frame, and interpretation of determination decided for that context. Thus there will be no formal specification of a developmental model, as such a structure bears no relationship to a real structure in nature, and in fact is alien to the phenomenon of locally mediated change. Rather the "developmental model" is the aggregate of all its submodels acting together fortuitously. Particular developmental models must therefore be presented informally as such aggregates.

University of Alberta

6.5.1 Modelling the Blastoderm of Drosophila

Kauffman et al. (1978) have characterized the blastodermal stage of Drosophila in terms of a code of binary switches constructed in response to positive and negative variations from the uniform state of a reaction-diffusion system. His model agrees with experimental

findings in predicting initial "segmentation" of the elongated, ellipsoidal egg, as sinusoidal waves fit into its length, followed as its circumference enters the first critical size, by division of the egg into dorsal and ventral compartments. Furthermore the codes thus assigned agree remarkably well with transdetermination data taken from experiments such as Hadorn's (1966), in that the more digits that differ, the lower the transdetermination rate.

Implicit in Kauffman's model are certain assumptions concerning cell division in the egg. The size of the egg does not increase, but nucleation alters the reaction-diffusion parameters, by decreasing diffusion rates, in a manner which allows successive patterns to be supported in the egg as their wavelengths decrease. The rate of nucleation is uniform throughout the egg, and this rate is such that odd patterns after the first (patterns asymmetric with respect to anterior end of the egg) are missed.

A model has been constructed in which these assumptions are formalized.

In the growth model synchronous, autonomous cell division is selected because it (a) provides uniform expansion, (b) results in odd modes beyond the first being missed, and (c) is mathematically equivalent to variation of system parameters as described above (Kauffman et al. (1978)).

The growth time step and communication time step are the same. Furthermore, while a general background of

University of Alberta

unoriented cell divisions is assumed, the orientation of divisions down the egg, due to flows during pattern formation, is modelled in order to maintain an elongated shape.

Since the shape functions on ellipses are not yet available, the blastoderm is modelled here as a rectangular structure. A code fragment "10" is added to a cell's genetic code whenever a pattern arises in which species one is positive and species two negative, and "01" when species one is negative and species two positive. For purposes of display these codes have been translated into letters as follows:

A - 10	I - 100110
B - 01	J - 100101
C - 1010	K - 011010
D - 1001	L - 011001
E - 0110	M - 010110
F - 0101	N - 010101
G - 101010	O - 010100
H - 101001	P - 010011

These codes translate directly to Kauffman's setting of '1' for species one positive, '0' for species one negative.

The results of simulation may be seen in Figure 6.18. Note that the width of the system is noncritical.

The assignment of codes coincides (under translation) with Kauffman's and therefore agrees with the transdetermination data. Segments are observable, though a

University of Alberta

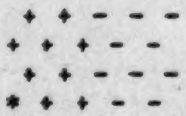


FIGURE 6.18a Segmentation in the blastoderm,

nodal pattern one

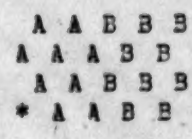


FIGURE 6.18b Assignment of code

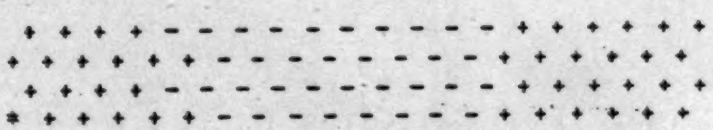


FIGURE 6.18c Nodal pattern two

```
 C C C C D D D D F F F F F E E E E E E
C C C C C C D D D D D F F E E E E E E
 C C C C D D D D F F F F F E E E E E E
* C C C C C D D D D D D F F E E E E E E
```

FIGURE 6.18d Assignment of code

FIGURE 6.18e Nodal pattern three

```
GGGG
```

mixture of two cell types is seen near the center of the structure. This intermixing of cells is a result of the discrete representation in hexagonal coordinates and is caused by decisions dependent on vertical lines being imposed on the array. Any discrete representation must have limitations in some circumstances, however. The results for the blastoderm, in showing that segments do indeed arise as determination occurs in response to reaction-diffusion in an elongated shape, are judged to be largely successful.

6.5.2 Modelling the Chick Wing

The chick wing is a growing structure in which cell differentiation into bone is observable. Initially a limb bud appears and begins to grow outwards. During the elongation phase, cells in the proximal central portion of the wing are observed "condensing" into bone tissue; this structure follows the process of growth down the limb. At the "elbow" the bone structure becomes double, and at the wrist it separates into three digits.

Many models have been proposed for the regulation of these events. Wolpert's (1978) introduces the concept of the progress zone in which cells can "count" the time they spend in the distal portion of the wing, and after emerging, use this information to differentiate. Newman and Frisch (1979) suggest a distal diffusion chamber in which the pattern of cartilage formation in the plane perpendicular to the proximo-distal axis of the wing is specified by critical

University of Alberta

patterns in a reaction-diffusion system. Newman and Frisch suggest another mechanism for proximo-distal pattern formation. Wilby and Ede (1975) have developed a simulation program in which artificially imposed periodic functions are used as prepatterns for bone formation. Moderate success has been achieved in obtaining realistic-looking patterns in a rectangular coordinate system. Wilby and Ede (1975) do not here model growth or intercellular processes explicitly.

Newman and Frisch's (1979) model with a distal diffusion chamber has two flaws. Firstly, the cross-section is modelled as rectangular, thus being unable to produce an isolated central area of bone tissue. Secondly, the model has to call upon a second mechanism to explain the proximo-distal bone patterns.

A model is presented here which can account for all the features of the phenomenon and can be observed in action via simulation. It is described briefly below:

(1) Assume a background rate of general uniform expansion of the system.

(2) The initial configuration in the limb bud is an elliptical cross-section perpendicular to the proximo-distal axis which is of a size critical for a reaction-diffusion system such as the Kauffman system. The shape function is given by the first Matthieu function, yielding a central elliptical core of values negative in species one and positive in species two, surrounded by cells with opposite signs. This configuration may be seen in Figure 6.19. The

University of Alberta

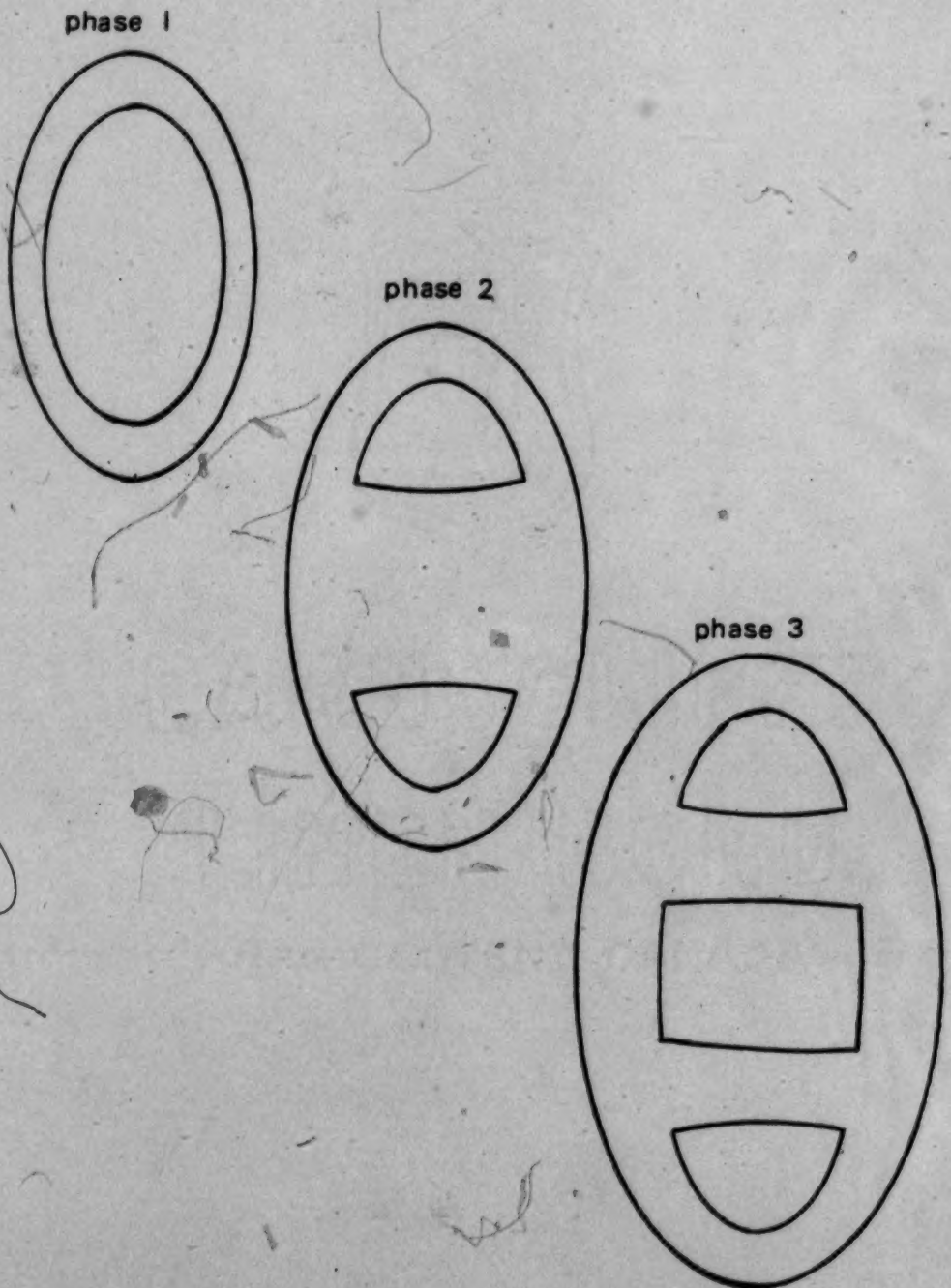


FIGURE 6.19 Elliptic cross-sections in the wing

cross section perpendicular to the dorso-ventral axis is assumed to be rectangular and is noncritical in the proximo-distal direction.

(3) During general expansion, the elliptical pattern is not destroyed, but the proximo-distal dimension reaches critical size for pattern 1 of a mixed boundary value problem. The proximal boundary of the limb (continuous with body tissues) does not obey a zero normal flux condition but rather conforms to noncritical values, that is, to the uniform stable state.

(4) In the proximal half of the limb bud, the positive contribution from the new pattern reinforces the existing pattern caused by the elliptic cross-section, while in the distal portion the patterns are reversed in sign.

(5) The reinforcement of the pattern means that proximal tissue is held at concentration values away from uniformity for a prolonged length of time. Through either deprivation (causing repression) or saturation (causing derepression) this results in determination and differentiation of these cells into bone cells, in the central core, and other tissue, in the surrounding cells.

(6) The differentiated cells are no longer capable of supporting the reaction-diffusion system, so the system's inner boundary moves distally.

(7) The system is no longer in critical size in the proximo-distal direction, so it returns to uniformity in that direction though still supporting the elliptical cross-

University of Alberta

section. In returning to uniformity it causes cell divisions to be oriented down the limb, thus causing it to elongate.

(8) The proximo-distal dimension again becomes critical for pattern 1, and steps (4)-(8) repeat, until

(9) Owing to the background growth, the ellipse leaves the critical range for the central core pattern and enters the critical range for another pattern in which along the major axis of the ellipse (anterior-posterior axis) there are two convex negative regions for species one, while the rest of the ellipse is positive (one of the Matthieu functions). Steps (3)-(8) repeat until

(10) During general expansion, the ellipse reaches a third critical phase in which three negative areas occur on the major axis of the ellipse (another Matthieu function). Steps (3)-(8) repeat until growth is stopped.

The model has a "diffusion chamber" or "progress zone" of undifferentiated cells at the distal end of the limb. Surgical experiments on chick wings in which the distal portion is reduced have resulted in malformed, randomly differentiated limbs. The model would predict such a result because elongated shape depends on pattern establishment at certain sizes and differentiation depends on pattern 1 occuring at these sizes. So reducing the size destroys these mechanisms. The model is sufficient throughout the development of the wing; that is, all aspects of the shape and differentiation phenomenon can be explained by this one

University of Alberta

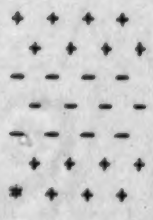
mechanism.

There is an analogy to Wolpert's clocked time in the progress zone, since differentiation is provoked by a certain "extended time" in a situation of deprivation or saturation. It is felt however that the present model is much closer to the spirit of locally mediated changes. Furthermore it explains geometry as well as differentiation. Limbs exhibit elliptic cross-sections, with more eccentric ellipses enclosing more bone structures in a proximo-distal succession.

Prolonged deprivation or saturation as a chemical basis for differentiation may have significance in terms of the portions of DNA in which repeated sequences occur. For example, a state of prolonged deprivation of a substance, which is normally bonded to such an area in multiple copies, might result in evacuation of the substance to maintain the standing pattern. Removal of all such bonded copies, after the pattern has been in place for a critical time period, might result in the derepression of portions of DNA coding for differentiation into cartilage. Such an interpretation agrees with the identification of repeated sequences as possible timing mechanisms, which was mentioned in Chapter Three.

The simulation of the chick wing model may be seen in Figures 6.20-6.22. Cells marked by "Z" are undetermined, while "B" marks cartilage cells and "A" the surrounding tissue. The three phases of growth correspond to the three

University of Alberta



SCALE 1.000000 TO 1
28 CELLS AFTER 0 GROWTH CYCLES

FIGURE 6.20a Limb bud, nodal pattern, phase one

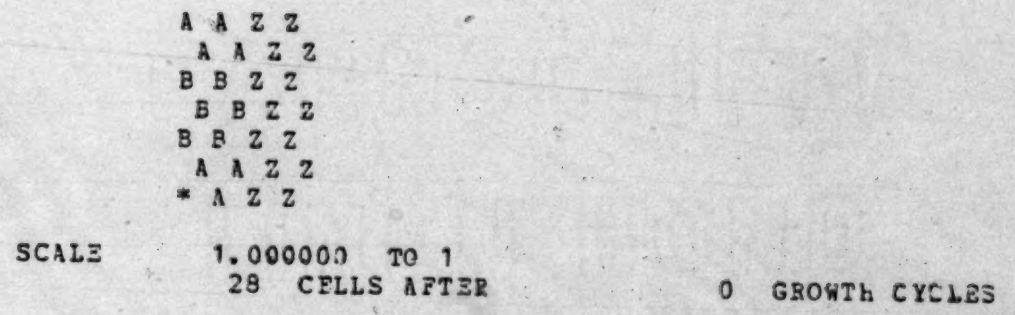
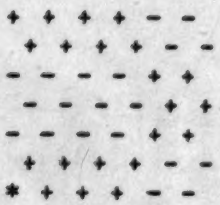


FIGURE 6.20b Limb bud, cell determination, phase one

```
A A Z Z Z Z
 A A Z Z Z Z
B B Z Z Z Z
 B B Z Z Z Z
B B Z Z Z Z
 A A Z Z Z Z
* A Z Z Z Z

SCALE      1.000000  TO 1
             42  CELLS AFTER          1  GROWTH CYCLES
```

FIGURE 6.20c Growth, phase one



SCALE 1.000000 TO 1
42 CELLS AFTER 1 GROWTH CYCLES

FIGURE 6.20d Nodal pattern, phase one

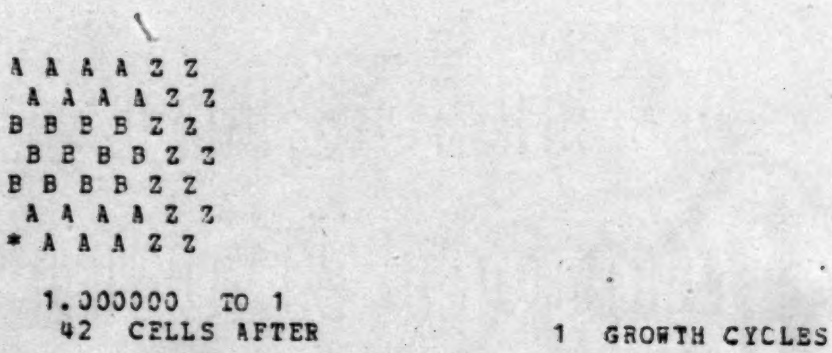
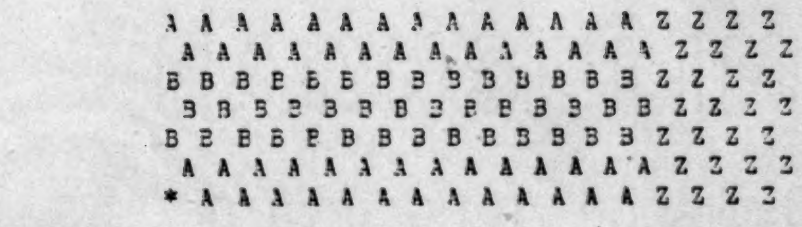


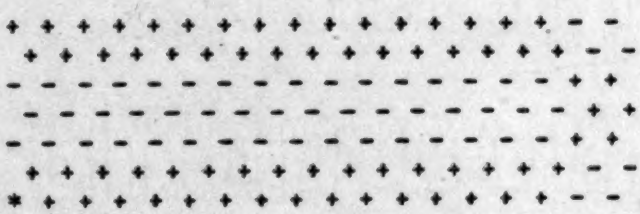
FIGURE 6.20e Cell determination, phase one



SCALE 1.000000 TO 1

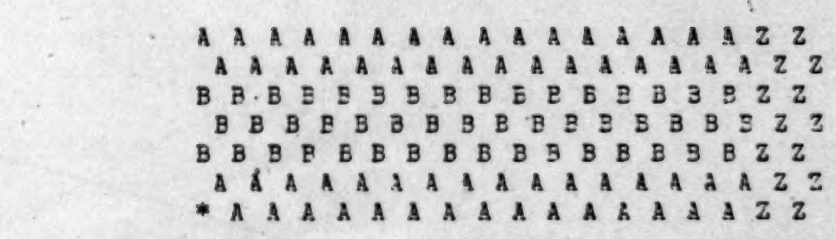
126 CELLS AFTER 7 GROWTH CYCLES

FIGURE 6.20f Growth, phase one



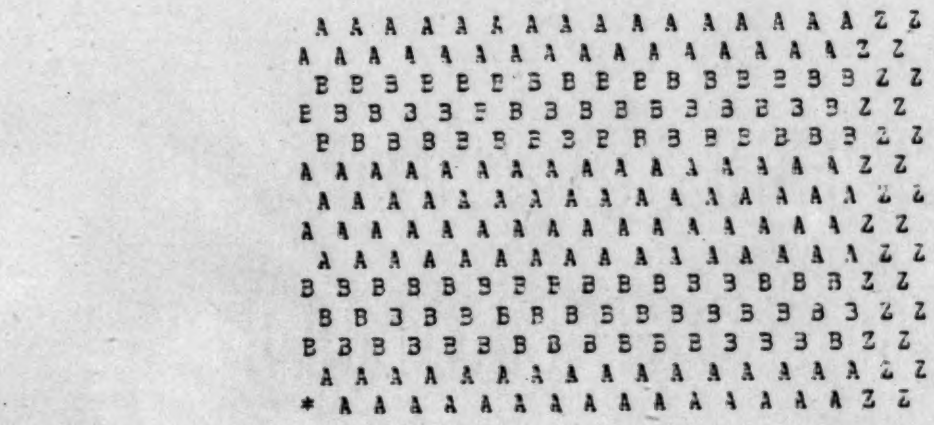
SCALE 1.000000 TO 1
126 CELLS AFTER 7 GROWTH CYCLES

FIGURE 6.20g Nodal pattern, phase one



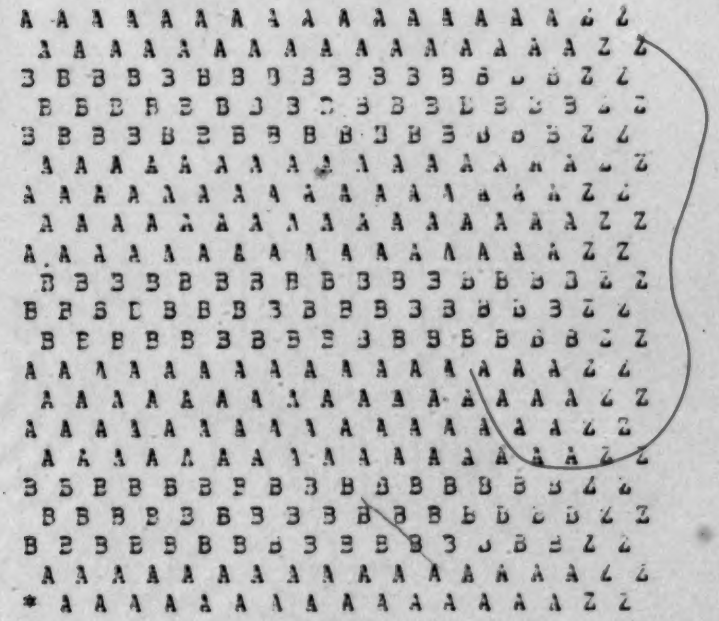
SCALE 1.000000 TO 1
126 CELLS AFTER 7 GROWTH CYCLES

FIGURE 6.20h Cell determination, phase one



SCALE 1.000000 TO 1
252 CELLS AFTER 7 GROWTH CYCLES

FIGURE 6.21 Cell determination, phase two



SCALE 1.000000 TO 1
378 CELLS AFTER 7 GROWTH CYCLES

FIGURE 6.22 Cell determination, phase three

elliptic cross sections of Figure 6.19.

6.6 Conclusion

The development of a framework for modelling and simulating development is described in this chapter. Models of cellular and intercellular processes are simulated, and conditions under which they may be combined in modelling developmental systems are discussed. These processes include cell division, the control of growth in cell populations, intercellular communication, pattern formation, determination and differentiation. Within the modelling structure, various options for component models are provided. The framework is constructed so that at each level different components for new phenomena, or atlternatives for existing models, could be incorporated. Application of the modelling and simulation framework to two developmental systems, the blastoderm of Drosophila and the chick wing, is described.

University of Alberta

CHAPTER SEVEN

Conclusions

University of Alberta

The primary goal of this project, as stated in Chapter One, is to provide a functioning and useful framework for modelling in developmental genetics. Along with consideration of success in achieving this goal, contributions to the art of modelling and simulation and to the establishment of a mathematical context for development can be discussed. Finally, Ziegler's theory of modelling and simulation can be assessed in practical application to this type of modelling situation.

Phenomena at various levels in developing systems have been modelled and simulated successfully. The modelling framework includes (1) models of cell division and growth with locally mediated control mechanisms, (2) models, both numeric and analytic, of reaction and diffusion as mechanisms of pattern formation, and (3) models of

determination and differentiation. Two developmental systems, employing a combination of these modelling options, are presented. These are segmentation in the blastoderm of Drosophila and bone condensation and growth in the chick wing.

Some observations should be made about the contributions of these developmental models. Firstly, the various phenomena and processes of development have been modelled and simulated in interaction, as is possible only through a framework in which different models can be combined. Secondly, all models in the framework adhere strictly to the requirement that changes be locally mediated. Thus, for example, departures from initially uniform concentration configurations are not induced by an imposed polarization of some kind, but by a local process (reaction-diffusion) considered in an intercellular context (the morphogenetic field). Similarly, cells experience genetic changes individually in response to critical values of local influences. Fundamental to the ability of the framework to model interacting processes, and locally mediated change, is the provision of modelling structures which allow sufficient information to be maintained for each cell, and permit modelling to be done at higher levels of the developmental system.

At each level of modelling there are opportunities for further options to be developed. Different models of growth, both intracellular and intercellular, could be

University of Alberta

formulated. Intercellular communication other than by reaction and diffusion could be studied (if other stable mechanisms were found). Determination and differentiation in a variety of one and two dimensional systems could be modelled; for example, experiments in the ligation of insect eggs could be performed (Kauffman et al., 1978). It is a measure of the power and flexibility of the modelling framework that such extensions are immediately feasible.

The modelling framework is significantly limited, in its present state of development, by difficulties in studying the establishment of patterns in non-rectangular domains. On the one hand, numerically integrated solutions in such domains are expensive (though feasible) to obtain. On the other hand, the shape functions with which analytic solutions can be calulated may be unknown (as in irregular domains) or difficult to use (as in circular or elliptic domains).

It may be noted, however, that if numerical integration methods are used, they are expensive only during critical phases, when the system is either leaving or returning to the uniform, steady state configuration. During those periods of time when the system is noncritical, convergence to the steady state solution is immediate at each time step. Furthermore patterns such as those made up of Bessel and Mathieu functions may be obtainable analytically by using numerical techniques to generate polynomial approximations in their respective domains. The scope of the modelling

University of Alberta

framework would be increased considerably by the addition of such techniques. Modelling embryonic development, imaginal discs in Drosophila, and other systems decomposible into (approximately) regular domains, would then become feasible in the framework.

The data structures of the modelling framework could be employed fruitfully in other biological simulations where the capacity for dynamic, locally mediated changes in cell space structures is required. Such areas include studies of cell movements, cell sorting models, and brain models. It may be noted with regard to the last that reaction-diffusion systems are being investigated in modelling memory (Tuckwell, 1979). Experiments with cancer and clonal studies of mutated cells can also be studied in this framework.

University of Alberta

The framework is thus able to accommodate a resonable variety of phenomena in the biological domain. This capacity is felt to be an indication of structural validity in the many component models of the framework and in the ways in which they can be combined.

Simulation of the models in the framework makes special contributions to developmental modelling. Firstly, in order for behavioral data to be generated by simulation, every aspect of that simulation, and of the models it represents, must actually perform as expected operationally. Thus in the model of Section 6.5.2 pre-patterning and differentiation in the chick wing can be observed in action.

A faulty model such as that of Newman and Frisch (1979) could not generate such behavior. In contrast, the model of Kauffman et al. (1978) of the blastoderm does generate its predicted behavior. Simulation, when performed with regard to validity of the iterative representation of a model, fulfils the function of ensuring that the model is a valid generator of the desired behavior. Secondly, through simulation an inferred process known as "reaction and diffusion" has been demonstrated to be a viable tool in pattern formation. Thirdly, the vital importance of the relative speed with which growth takes place and reaction-diffusion operates has been illustrated by means of simulation. Which patterns appear in a system, and therefore what structures of differentiated and determined cells will occur, depend on this timing. It is suggested that the interaction of these phenomena (growth and communication), in combination with viability tests in the process of natural selection, may play a part in the evolution of species.

In view of the above contributions, it would appear that a useful, functioning framework for developmental modelling has been largely achieved. Contributions to the art of modelling and simulation may now be discussed.

The modelling framework is multilevelled: modelling is possible at the molecular level (reaction and diffusion of chemical species, for example), at the cellular level (cell division, for example), and at the level of the

University of Alberta

developmental system (the chick wing, for example). The framework is also of mixed type: state variables may be discrete (clocks), binary digit strings (genetic code), or real numbers, (values in a continuous field such as local concentrations). The framework forms a "model pool" of the type envisaged by Ziegler (1976), and it has been demonstrated how selected members of the pool act together in obedience to conditions of validity established with respect to their common time base.

The modelling framework, therefore, is a functional example of aspects of modelling and simulation of current interest. The techniques developed to deal with dynamic change in cell structures and locally mediated processes (local transition functions) may prove valuable in any area of computing science where the cell space formalism is used. In summary, it is hoped that contributions have been made to the art of modelling and simulation.

The mathematical context which has been presented for development is a systems theoretic formalism operating in a cell space. Through its components (cells and lists) it proves to be sufficient in representing the structures of the chosen experimental frame. Through its transition functions the formalism appears capable of representing dynamic processes in these structures. Such processes employ those mathematical tools, pertaining to various levels of the system, which are identified in Chapter One as being necessary components of developmental modelling. For

University of Alberta

example, differential equations are employed at the molecular level, statistical techniques are used in modelling cell growth in a population, and equilibrium phenomena are employed in morphogenetic fields.

The formalism proves to be capable of representing dynamic changes in structures, that is in representing the interaction of the dynamic and static elements of the modelling domain. An example of change in static elements or structure is given by growth, which may be controlled by dynamically changing, locally mediated processes. An example of change in dynamic elements is given by determination, in which the genetic code changes in response to critical values of some local phenomena, resulting in changed activity.

Simulation of the system specifications forms a vital part of the mathematical context which has been employed. Dynamic aspects of models without simulations can only be appreciated by means of inference from the transition function specifications, whereas simulation actually generates behavior for the model. The transition functions can thus be observed in action, and since they are meant to describe action this is the appropriate technique. Questions of validity must of course be addressed; this may be done directly since Ziegler's (1976) theory of modelling and simulation operates in the same formalism. In many respects, therefore, the mathematical context chosen for modelling the developmental system proves to be

University of Alberta

satisfactory.

Ziegler's theory of modelling and simulation (Ziegler, 1976) has been used throughout in the development of the modelling framework and in its exposition. The formal specifications of models in Chapter Six proved to be cumbersome at times, yet this is justifiable on two counts. Firstly, sets of component models of the "modelling pool" can be selected with regard to formally established criteria for validity (as, for example, in the selection of a growth and signalling model pair). Secondly, changes in structure and local functioning are describable in systems theoretic terms in a manner which relates them to global structure; this perspective could not be gained if they were to be described simply as individual algorithms.

Morphisms between models at the iterative level have not proved useful. It is felt that they might become applicable if the modelling framework were to be extended. The primary thrust of a project such as this is in establishing the different levels of the modelling framework (a "vertical" phase of development). Later expansion, by means of other models, and lumped models of existing models, would occur at each level (a "horizontal" phase of development). Relationships between models on a particular level could then be investigated, by means of canonical decompositions of input segments and the relevant morphisms.

Ziegler's theory has proved invaluable as a means of keeping perspective in the construction of a modelling

University of Alberta

framework of this complexity and scope. Organized progress from the real system to simulation experiments, as followed in the thesis and in the development of the modelling framework, with attention paid to questions of validity at each stage, has resulted in a functioning, effective modelling framework.

In summary, it may be stated that many of the goals of the project have been accomplished. A useful framework for modelling and simulation in developmental genetics has been accomplished, with, it may be hoped, benefit to both biology and computing science.

University of Alberta

REFERENCES

Alberghina, F.A.M. A model for the regulation of growth in mammalian cells. Journal of Theoretical Biology 55, 533-545, 1975.

Alberghina, F.A.M. Dynamics of the cell cycle in mammalian cells. Journal of Theoretical Biology 69, 633-643, 1977.

Alberghina, F.A.M. Modelling the control of cell growth. Simulation 31, 37-41, 1978.

Babloyantz, A., and Hiernaux, J. Models for positional information and positional differentiation. Proceedings of the National Academy of Sciences 71, 1530-1544, 1974.

Babloyantz, A. Mathematical models for morphogenesis. Mathematical Biology 37, 637-669, 1975.

Baker, R.W., and Hermann, G.T. Simulation of organisms using a developmental model, Part 1: Basic description, Part 2: The heterocyst formation in blue-green algae. Bio-Medical Computing 3, 201-215, 251-267, 1971.

Britten, R.J., and Davidson, E.H. Gene regulation for higher cells: A theory. Science 165, 349-357, 1969.

Bryant, P.J. Regeneration and duplication following operations in situ on the imaginal discs of Drosophila melanogaster. Developmental Biology 26, 637-651, 1971.

Child, C.M. Patterns and Problems of Development. University of Chicago, Chicago, 1940.

Crick, F.H.C. Diffusion in embryogenesis. Nature 225, 420-427, 1970.

Crick, F.H.C., and Lawrence, P.A. Compartments and polyclones in insect development. Science 189, 340-347, 1975.

Duchting, W. Regulation. (manuscript), 1977.

Ericson, R. Growth in two dimensions, descriptive and theoretical studies, in Automata, Languages, Development, eds. Lindenmayer, A. and Rozenberg, G., North-Holland, Amsterdam, 1976.

Fisher, R.A. Genetical Theory of Natural Selection. Oxford University Press, Inc., New York, 1930.

Fishman, G.S. Concepts and Methods in Discrete Event Simulation. Wiley-Interscience, New York, 1973.

University of Alberta

Fox, L. Numerical Solution of Ordinary and Partial Differential Equations. Pergamon Press, New York, 1962.

French, V., Bryant, P.J., and Bryant, S.V. Pattern regulation in epimorphic fields, Science 193, 969-981, 1976.

Garcia-Bellido, A. Genetic control of development, in Cell Patterning, CIBA Foundation Symposium 29, Associated Scientific Publishers, Amsterdam, 1975.

Garcia-Bellido, A., Ripoll, P., and Morata, G. Developmental compartmentalisation of the wing disc of Drosophila. Nature New Biology 245, 251-253, 1973.

Gierer, A., and Meinhardt, H. A theory of biological pattern formation. Kybernetik 12, 30-39, 1972.

Gmitro, J.L., and Scriven, L.E. A physicochemical basis for pattern and rhythm, in Intercellular Transport, ed. K.B. Warren, Academic Press, 221-255, 1966.

Goodwin, B.C. Mathematical metaphor in development. Nature 242, 208-223, 1973.

Grossberg, S. Communication, memory, and development, in Progress in Theoretical Biology, eds. Rosen, R., and Snell, F.M. Academic Press, 183-231, 1978.

Grossberg, S. Decisions, patterns and oscillations in nonlinear competitive systems with applications to Volterra- Lotka systems. Journal of Theoretical Biology 73, 101-130, 1978.

Hadorn, E. Problems of determination and transdetermination. Brookhaven Symposium 8, 148-159, 1965.

Haldane, J.B.S. Enzymes. Longmans, Green Company Ltd., London, 1930.

Haldane, J.B.S. The Causes of Evolution. Longmans, Green Company Ltd., London, 1932.

Herman, G.T. and Liu, W.H. The daughter of Celia, the French flag, and the firing squad. Simulation 28, 33-41, 1973.

Herman, G.T., and Rozenberg, G. Developmental Systems and Languages. North-Holland/American Elsevier, New York, 1975.

Herman, G.T., and Schiff, G.L. Simulation of organisms based on L systems, in Proc. of the 1974 Conference on Biologically Motivated Automata Theory, IEEE, New York, 70-76, 1974.

Hogeweg, P. Simulating the growth of cellular forms. Simulation 31, 90-96, 1978.

Kauffmann, S.A. Control circuits for determination and

University of Alberta

transdetermination. Science 181, 310-323, 1973.

Kauffmann, S.A., Shymko, R.M., and Trabert, K. Control of sequential compartment formation in Drosophila. Science 199, 259-270, 1978.

Kiefer, J. A model of feedback-controlled cell population. Journal of Theoretical Biology 21, 263-279, 1968.

Kim, M., Bahrami, K., Woo, K.B. Mathematical description and analysis of cell cycle kinetics and the application to Erlich Ascites tumor. Journal of Theoretical Biology 50, 437-459, 1975.

Lawrence, P.A. Compartment boundaries in Oncopeltus, in Cell Patterning, CIBA Foundation Symposium 29, Associated Scientific Publishers, Amsterdam, 1975.

Lawrence, P.A., Crick, F.H.C. and Munro, M. A gradient of positional information in an insect, Rhodnius. Journal of Cell Science 11, 815-853, 1972.

Lawrence, P.A. and Morata, G. The early development of mesothoracic compartments in Drosophila. Developmental Biology 56, 40-51, 1977.

Lewis, J.H. Rules for building the chick wing: discrete and continuous aspects of morphogenesis, in Automata, Languages, Development, ed. Lindenmayer, A., and Rozenberg, G., North Holland, Amsterdam, 25-38, 1976.

Lindemayer, A. L-systems in their biological context. Proceedings of the 1974 Conference on Biologically Motivated Automata Theory, IEEE, 65-70, 1974.

Lindenmayer, A., and Rozenberg, G. Automata, Languages, Development. North-Holland, Amsterdam, 1975.

MacWilliams, H.K., and Papageorgiou, S. A model of gradient morphogen binding. Journal of Theoretical Biology 72, 385-411, 1978.

Meinhart, H. A model of pattern formation in insect embryogenesis. Journal of Cell Science 23, 111-142, 1978.

Newmann, S.A., and Frisch, H.L. Dynamics of skeletal pattern formation in the developing chick limb. Science 205, 662-668, 1979.

Nicolis, G., and Prigogine, I. Self-Organization in Nonequilibrium Systems. Wiley-Interscience, New York, 1977.

Ransom, R. Computer analysis of cell division. Simulation 28, 189-197, 1976.

University of Alberta

Ransom, R. Computer analysis of cell division in Drosophila discs: model revision and extension to simulate leg disc growth. Journal of Theoretical Biology 66, 361-377, 1977.

Raven, C.P., and Bezem, J.J. Analysis of pattern formation in gastropods by means of computer simulation, in Automata, Languages, Development, ed. Lindenmayer, A., and Rozenberg, G., North Holland, Amsterdam, 147-159, 1976.

Raveshky, N. Mathematical Biophysics. Dover, New York, 1960.

Reiner, J.M. The Organism as an Adaptive Control System. Prentice-Hall, Academic Press, 1972.

Rosen, R. Foundations of Mathematical Biology, Vol. II, Cellular Systems. Academic Press, 1972.

Sampson, J.R. Biological Information Processing: Current Theory and Computer Simulation. In preparation, 1980.

Schubiger, G. Regeneration, duplication and transdetermination in fragments of the leg disc of Drosophila melanogaster. Developmental Biology 26, 277-295, 1971.

Thom, R. Stabilite structurelle et morphogenese: Essai d'une theorie generale de modeles, ed. Wightman, A.S., Addison-Wesley, New York, 1972.

Tuckwell, H.C. Solitons in a reaction-diffusion equation. Science 205, 493-501, 1979.

Turing, A.M. The chemical basis of morphogenesis. Philosophical Transactions of the Royal Society of London B 237, 37-56, 1952.

Waddington, C. H. Organisers and Genes. Cambridge University Press, London, 1940.

Weischaus, E., and Gehring, W. Gynandromorph analysis of the thoracic disc primordia in drosophila melanagaster. William Roux Archives 180, 31-46, 1976.

Wheldon, T.E., Kirk, J., and Gray, W.m. Mitotic autoregulation, growth control and neoplasia. Journal of Theoretical Biology 38, 627-639, 1973.

Wilby, O.K. and Ede, D.A. A model generating the pattern of skeletal elements in the embryonic chick wing, in Proceedings of the 1974 Conference on Biologically Motivated Automata Theory, IEEE, 81-90, 1974.

Wilby, O.K., and Ede, D.H. Computer simulation of vertebrate limb development - the effect of cell division control patterns, in Automata, Languages, Development, ed. Lindenmayer, A., and Rozenberg, G., North Holland,

University of Alberta

Amsterdam, 25-38, 1976,

Wilby, O.K., and Webster, G. Transmission of hypostome inhibition. Journal of Embryology and Experimental Morphology 24, 583-615, 1970.

Wolpert, L. Positional information and the spatial pattern of cellular differentiation. Journal of Theoretical Biology 25, 1-47, 1969.

Wolpert, L. Positional information and pattern formation. Current Topics in Developmental Biology 6, 183-221, 1971.

Wolpert, L. Pattern formation in biological development. Scientific American 239, 154-164, 1978.

Wright, S., Evolution in Mendelian populations. Genetics 16, 97-159, 1931.

Zahler, R.S., and Sussman, H.J. Claims and accomplishments of applied catastrophe theory. Nature 269, 759-767, 1977.

Zeigler, B.P. Theory of Modelling and Simulation. Wiley-Interscience, New York, 1976.

University of Alberta